## Supporting Information

*From Conquest to Centralization: Domestic Conflict and the Transition to Direct Rule*

Following text to be published online.

## Contents


**A Model** **2**

A.1 Domestic Conflict and the Transition to Direct Rule . . . . . 2

A.2 General Form for the Probability of Successful Potentate Resistance, $\rho(\cdot)$ . . . . . 9

A.3 Marginal Changes in Population . . . . . 11

**B Data Appendix** **12**

B.1 Descriptive Statistics . . . . . 12

B.2 Balance on Observables in 1550 . . . . . 14

B.3 Geographic Distribution of Drought-Rain Shocks . . . . . 15

**C Additional Evidence** **16**

C.1 Reduced-Form Results . . . . . 16

C.2 Empirical Analysis including Singleton Observations . . . . . 17

C.3 Empirical Analysis with Tributaries . . . . . 19

C.4 Empirical Analysis with 10-Year Window for Cutoffs . . . . . 23

C.5 Empirical Analysis with Balanced Panel . . . . . 27

C.6 Weak-Instrument Robust Inference . . . . . 28

C.7 Alternative Instrument Construction . . . . . 30

**D Additional Evidence: Mechanisms** **32**

**E Additional Evidence: Persistent Effects of Early Direct Rule** **40**

**F Appendix References** **41**

# A. Model

## A.1 Domestic Conflict and the Transition to Direct Rule

We present a two-period model that considers the strategic interaction between a ruler and a local potentate. The model begins under indirect rule, where the potentate maintains order in the region in exchange for a share of tax revenue. We show that if both tax revenue and the risk of rebellion are increasing in local population, as we argue was the case in early colonial Mexico and in other settings, a precipitous decline in population will lead to a higher likelihood of a transition to direct rule. Furthermore, we show that the effect of the population decline is amplified in regions where collective action and rebellion are easier (as this raises the value of indirect rule) and where elites have more profitable outside options (which decreases elite incentives to contest the centralization of fiscal authority).

**Actors and timing:** Consider two actors, a ruler ($R$) and a representative local elite ($E$), who interact in a two period game, $s = \{1,2\}$. In each period, the potentate raises taxes from the subjects in his district, $T$. He keeps an exogenous portion $\gamma \in (0,1)$, and transfers the rest to the ruler. The local population, in response to extractive taxation, can rebel in the second period with an exogenous probability $W \in [0,1]$. This probability is an increasing function of the total number of subjects ($T$) and the ease of local collective action ($\alpha$): $p(W=1) = \omega(T,\alpha)$, with $\omega_\alpha(T,\alpha) > 0$ and $\omega_T(T,\alpha) > 0$. The assumption that the probability of rebellion is increasing in $T$ is supported by both Mexico-specific literature on unrest in the colonial period (e.g., Taylor, 1979; Katz, 1988) and by a broader literature linking population density with conflict (e.g., Boserup, 1965; Homer-Dixon, 1999; Goldstone, 2002). Some factors that may facilitate local collective action include sharing a common language, which can enhance the population's ability to coordinate.

In the first period, the ruler can choose to implement direct rule (decision $D$) at a direct administrative cost $C_D$, and strip the potentate of his formal rights. If successful, this initiative leads to direct rule in the second period, which allows the ruler to capture all of the revenue in the region,

$T$, without losing a fraction $\gamma$ to the potentate. We normalize the tax revenue to 1 unit per local taxpayer.

The local potentate, in turn, makes two decisions in period one. First, he decides whether to pay a cost, $C_G$, to guard his region against rebellion (decision $G$). Second, after observing any attempt by the ruler to establish direct rule, the potentate can select a share of his first-period income to resist the ruler's efforts to centralize power, $r \in [0,1]$. The potentate's income is a function of the local tax-paying population, $\gamma T$. The probability that an attempt to displace the potentate fails is given by the concave function $\rho(r\gamma T) \in [0,1]$, which is increasing in the intensity of his resistance against the ruler, $r\gamma T$. For simplicity, we assume a specific functional form, $\rho(r\gamma T) = \frac{\sqrt{r\gamma T}}{\bar{T}}$, where $\bar{T}$ is a finite upper bound for population in the region. In Appendix Section A.2, we characterize the results using a general form for $\rho(\cdot)$.

In period 1, the ruler receives:

$$U_1^R = \begin{cases} (1-\gamma)T & \text{if } D = 0 \\ (1-\gamma)T - C_A & \text{if } D = 1, \end{cases}$$

where $T$ are the capitation taxes levied from the population, and $C_A$ is the administrative investment cost necessary to attempt to establish direct rule in the second period.

Second-period payoffs depend critically on the probability of local rebellion. Whether this rebellion occurs depends on whether the potentate has decided to guard his region. For simplicity, we assume that the probability of rebellion is greater than zero when there is no protection, and zero when the region is protected by the potentate. In the second period, then, the ruler's expected payoff is:

$$E(U_2^R) = \begin{cases} (1-\gamma)T - \omega(T,\alpha)C_R & \text{if } D = 0 \text{ and } G = 0 \\ (1-\gamma)T & \text{if } D = 0 \text{ and } G = 1 \\ \left[1-\rho(r\gamma T)\right](T - \omega(T,\alpha)C_R) + \rho(r\gamma T)\left[(1-\gamma)T - \omega(T,\alpha)C_R\right] & \text{if } D = 1 \text{ and } G = 0 \\ \left[1-\rho(r\gamma T)\right](T - \omega(T,\alpha)C_R) + \rho(r\gamma T)\left[(1-\gamma)T\right] & \text{if } D = 1 \text{ and } G = 1, \end{cases}$$

where $\omega(T,\alpha)$ is the probability of rebellion in the second period in the absence of potentate protection, and $C_R > 0$ is an exogenous cost to putting down the rebellion.

The ruler can receive a higher share of the region's revenue if he successfully removes the potentate and sets up direct rule. This replacement, however, leaves the region unprotected against potential rebellion, captured by the non-zero probability $\omega(T,\alpha)$. An attempt to establish direct rule (i.e., $D=1$) can also prompt a reaction from the potentate, who can choose to resist the ruler's initiative by setting $r>0$. Resistance can improve the potentate's chances to keep his indirect-rule rights, which happens with probability $\rho(r\gamma T)$. On the other hand, if the ruler does not try to establish direct rule ($D=0$), he keeps receiving a smaller share the taxes, $(1-\gamma)T$. If, additionally, the potentate decides to set up a costly defense for the region ($G=1$), the risk of rebellion disappears. The tradeoff for the ruler is clear: the possibility of higher future revenue comes with an increased risk of rebellion.

For the local potentate, the payoff in the first period is:

$$U_1^E = \begin{cases} (1-r)\gamma T & \text{if } G=0 \\ (1-r)\gamma T - C_D & \text{if } G=1, \end{cases}$$

where $r$ is the share of his first-period income devoted to resist any attempt by the ruler to remove him, and $C_D$ is the cost of preparing the defense of the region against rebellion in the next period. The potentate's expected payoff in the second period is:

$$E(U_2^E) = \begin{cases} (1-\omega(T,\alpha))\gamma T & \text{if } D=0 \text{ and } G=0 \\ \gamma T & \text{if } D=0 \text{ and } G=1 \\ \rho(r\gamma T)(1-\omega(T,\alpha))\gamma T + [1-\rho(r\gamma T)]\underline{\mathrm{u}} & \text{if } D=1 \text{ and } G=0 \\ \rho(r\gamma T)\gamma T + [1-\rho(r\gamma T)]\underline{\mathrm{u}} & \text{if } D=1 \text{ and } G=1, \end{cases}$$

where $\underline{\mathrm{u}}$ is the potentate's outside option if direct rule is successfully implemented. The potentate only gets his share of tax revenue in the second period if no rebellion breaks out, and if the ruler decides not to establish direct rule (or the attempt is successful resisted.)

In short, the ruler decides whether to attempt to establish direct rule, while the potentate makes two choices in response: first, whether to resist the ruler's attempt, and whether to set up a defense against rebellion. The timing of the game is:

1. Parameters are given, first-period incomes are realized.

2. The ruler decides whether to attempt to establish direct rule (decision $D$)
3. The local potentate chooses the share of income used to resist direct rule, $r$, and decides whether to guard the region against rebellion (decision $G$).
4. If the ruler tried to establish direct rule, his attempt fails with probability $\rho(\cdot)$, and succeeds with probability $1-\rho(\cdot)$.
5. If the local potentate chose not to guard the region or if the ruler successfully establishes direct rule, rebellion breaks out with probability $\omega(T,\alpha)$. Second period incomes are realized.

**Solution.** We employ subgame perfection as a solution concept. We solve by backward induction, starting first with the local potentate's choice of whether to set up a costly defense against rebellion in his region (decision $G$). His decision simply weighs the benefits of protecting the region, given the risk of rebellion, against the cost of defending it. He chooses to defend if:

$$C_D \leq \begin{cases} \omega(T,\alpha)\gamma T & \text{if } D=0 \\ \rho(r\gamma T)\omega(T,\alpha)\gamma T & \text{if } D=1. \end{cases} \tag{A1}$$

That is, the potentate guards the region if the cost of defense is smaller than his expected share of future taxes, weighed by the risk of rebellion and the probability of successfully resisting any attempt by the ruler to remove him to set up direct rule. Note that, when the ruler does not attempt to establish direct rule (i.e., $D=0$), the potentate will have an incentive to protect his region even for higher defense costs.

Now we turn to the optimal choice of resistance, $r^*$, which is chosen simultaneously by the potentate. When the ruler decides not to establish direct rule, the potentate does not need to resist, and thus $r^*=0$. When his rights are challenged by the ruler, however, the potentate's optimal resistance is given by

$$r^* = \begin{cases} \left[\frac{1}{2\bar{T}}[(1-\omega(T,\alpha))\gamma T - \underline{\mathrm{u}}]\right]^2 & \text{if } G=0 \\ \left[\frac{1}{2\bar{T}}[\gamma T - \underline{\mathrm{u}}]\right]^2 & \text{if } G=1, \end{cases} \tag{A2}$$

which emerges from the potentate's utility maximization problem.[1]

[1] In Appendix Section A.2, we characterize the potentate's equilibrium behavior using a more general form for $\rho(\cdot)$.

The ruler, in turn, anticipates the potentate's actions and decides whether to establish direct rule. He attempts to implement direct rule if:

$$C_A \leq \begin{cases} \left[1-\rho(r_{G=0}\gamma T)\right]\left[\gamma T\right] & \text{if } G=0 \\ \left[1-\rho(r_{G=1}\gamma T)\right]\left[\gamma T-\omega(T,\alpha)C_R\right] & \text{if } G=1. \end{cases} \tag{A3}$$

In deciding whether to attempt to establish direct rule, the ruler weighs the costs and benefits given the expected reaction from the potentate. A successful establishment of direct rule enables the ruler to capture a higher share of the tax revenue. On the other hand, the costly attempt to set up direct rule can be sabotaged by the potentate with some probability, and, even if it succeeds, direct rule potentially exposes the region to rebellion.

**Comparative statics.** We now consider how a fall in population affects the establishment of direct rule in equilibrium. We focus on dramatic demographic shocks, such as the one experienced in the Americas following the Conquest. In Appendix Section A.3, we characterize how the equilibrium changes with smaller demographic shifts.

*Large shock to population*. Consider a decline in population that shifts the equilibrium from one in which condition (A1) is met and the local potentate decides to guard the region against rebellion to one in which this is no longer preferred. This discontinuity occurs because there is a population threshold, $\underline{T}$, below of which the potentate will no longer pay to defend his district as both the risk of rebellion and his own expected share of future tax revenue are declining in population. Because all of the terms on the right hand side of condition (A1) are declining in population—the risk of rebellion, the potentate share of tax revenue, and the likelihood of successfully resisting political centralization—this condition will necessarily be satisfied if the population declines enough.

Note that the probability of successful resistance to the attempted establishment of direct rule, $\rho(r^*\gamma T)$, includes not only the resources available to the potentate, $\gamma T$, but also the intensity of his opposition to direct rule, $r^*$. When the population threshold $\underline{T}$ is crossed and the potentate swiches from guarding his region ($G = 1$) to leaving it exposed to rebellion ($G = 0$), there is a discontinuous reduction in his equilibrium resistance to direct rule (i.e., $r^*_{G=1} > r^*_{G=0}$), as illustrated by condition (A2). This generates a discontinuity in the probability of successfully stopping the ruler's attempt to establish direct rule, $\rho(r^*_{G=1}\gamma T) > \rho(r^*_{G=0}\gamma T)$.[2]

Given this discontinuous change in the potentate's behavior, the ruler has a greater incentive to attempt to establish direct rule because of the increased likelihood that such an effort will succeed. This expands the range of administrative costs at which the ruler is willing to seek the establishment of direct rule. To see this directly, consider the maximum administrative cost that the ruler is willing to disburse to establish direct rule (condition (A3)) at population threshold $\underline{T}$. This cost is higher when the potentate stops defending his region if:

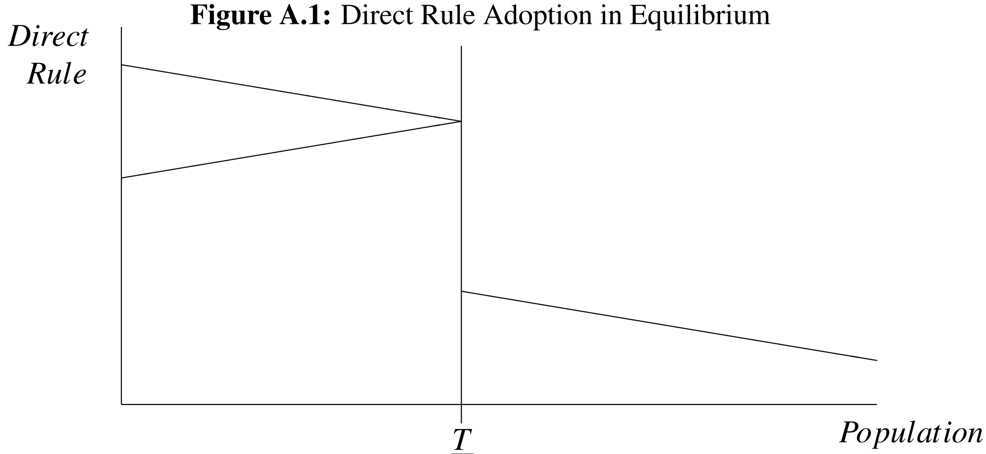
$$\left[1-\rho(r^*_{G=0}\gamma\underline{T})\right]\gamma\underline{T} \geq \left[1-\rho(r^*_{G=1}\gamma\underline{T})\right]\left[\gamma\underline{T}-\omega(\underline{T},\alpha)C_R\right],$$

which simplifies to

$$C_R \geq \frac{\gamma\underline{T}}{\omega(\underline{T},\alpha)}\left[\frac{\rho(r^*_{G=0}\gamma\underline{T})-\rho(r^*_{G=1}\gamma\underline{T})}{1-\rho(r^*_{G=1}\gamma\underline{T})}\right]. \tag{A4}$$

This condition is always met, and it implies that for a given administrative cost of establishing direct rule, the ruler is more likely to seek a transition to direct rule when the population threshold $\underline{T}$ is crossed.[3]

**Figure A.1:** Direct Rule Adoption in Equilibrium




[2]Specifically, optimal potentate resistance implies that $\rho(r^*\gamma T) = \frac{1}{2\bar{T}^2}[\gamma T - \underline{u}]$ if $G=1$, and $\rho(r^*\gamma T) = \frac{1}{2\bar{T}^2}[(1-\omega(T,\alpha))\gamma T - \underline{u}]$ if $G=0$. Thus, a discontinuous decline in potentate resistance from $r^*_{G=1}$ to $r^*_{G=0}$ leads to a discontinuous drop in the probability of successful resistance to direct rule, from $\rho(r^*_{G=1}\gamma T)$ to $\rho(r^*_{G=0}\gamma T)$.

[3]This is the case because the cost to put down a rebellion, $C_R$, is always positive by assumption, and the right hand side of the inequality must be negative (because $\rho(r^*_{G=1}\gamma T) > \rho(r^*_{G=0}\gamma T)$).

Figure A.1 illustrates the direct rule outcome in equilibrium as a function of population. Above the population threshold $\underline{T}$, marginal decreases in population lead to a higher likelihood of direct rule; below the threshold, the marginal effect of population is ambiguous (see Appendix Section A.3 for a characterization of these results). As population declines and crosses the threshold, there is a discontinuous jump in the likelihood of adopting direct rule.

Condition (A4) also suggests that the discontinuous jump in the maximum cost that the ruler is willing to spend in establishing direct rule is increasing in the baseline probability of rebellion, $\omega(\underline{T}, \alpha)$. To see this, note that the right hand side of condition (A4) declines with $\omega(\underline{T}, \alpha)$. For a given cost of putting down a rebellion, $C_R$, this implies a larger difference between the maximum administrative cost of direct rule, $c$, in condition (A3) when the potentate guards against rebellion, $G = 1$, as compared to when he does not, $G = 0$. Since $\omega(\cdot)$ is increasing in $\alpha$, if a region's characteristics facilitate collective action (i.e., the region has higher $\alpha$), the effect of the population collapse on the implementation of direct rule will be magnified.

*The value of the potentate's outside option.* We can draw one additional implication from the model related to the availability of an outside earnings option for the potentate. A more valuable outside option reduces the potentate's optimal level of resistance to any attempt by the ruler to establish direct rule. This is because being stripped of power becomes relatively less painful (see condition (A2)). Simultaneously, the potentate is less likely to defend the region for a given cost of defense (see condition (A1)). These changes in the potentate's behavior, make it more likely that the ruler will seek to establish direct rule. This is due both to the lower expected resistance from the potentate and the reduced likelihood that the potentate will provide an effective defense against local rebellion.

*Observable implications.* To summarize, we derive the following observable implications of the model:

(i) As population declines precipitously, the ruler is more likely to establish direct rule.

(ii) The effect of population collapse on the likelihood of direct rule is increasing in the rebelliousness of the region.

(iii) The effect of population collapse on the likelihood of direct rule is increasing in the value of the outside option available to the potentate.

(iv) More valuable outside options for the potentate lead to a higher likelihood of direct rule in equilibrium.

### A.2 General Form for the Probability of Successful Potentate Resistance, $\rho(\cdot)$

Leaving $\rho(\cdot)$ in a general form delivers the main results described in the text. As established above, when the ruler decides not to establish direct rule, the potentate will not resist (i.e., $r^* = 0$). When the ruler attempts to strip the potentate's rights, however, the optimal resistance is the result of the following optimization problem solved by the potentate:

$$max_{\{r\}} \begin{cases} (1-r)\gamma T + (1-\rho(r\gamma T))\underline{\mathrm{u}} + \left[\rho(r\gamma T)\right]\left[(1-\omega(T,\alpha))\gamma T\right] & \text{if } G=0 \\ (1-r)\gamma T + (1-\rho(r\gamma T))\underline{\mathrm{u}} + \left[\rho(r\gamma T)\right]\left[\gamma T\right] & \text{if } G=1. \end{cases}$$

We assume that $\rho(\cdot)$ is concave and increasing in $r$. Making no further functional form assumptions about $\rho(\cdot)$, the optimal level of resistance is implicitly given by the optimality condition:

$$\frac{\partial}{\partial r}\rho(r^*\gamma T) = \begin{cases} \dfrac{1}{(1-\omega(T,\alpha))\gamma T - \underline{\mathrm{u}}} & \text{if } G=0 \\ \dfrac{1}{\gamma T - \underline{\mathrm{u}}} & \text{if } G=1. \end{cases} \tag{A5}$$

This equilibrium behavior by the potentate is affected in the same way by a large shock to population when using a more general form for $\rho(\cdot)$. When the population threshold $\underline{T}$ is crossed, so that the potentate switches from guarding his region ($G = 1$) to leaving it exposed to rebellion ($G = 0$), there is a discontinuous reduction in his equilibrium resistance (i.e., $r^*_{G=1} > r^*_{G=0}$), which is visible in condition (A5). Note that as long as the potentate defends his region ($G = 1$) and the cost of putting down a rebellion are large enough (i.e., $C_R \geq \dfrac{\gamma}{\omega_T(T)}$), a marginal reduction in the population leads to an increase in the value of condition (A2). This implies that the level of resistance, $r^*$, which is implicitly defined in condition (A2), has to decline, since $\rho(\cdot)$ is a concave function of $r$. When the population threshold $\underline{T}$ is reached, the condition discontinuously increases in value, which implies a similarly discontinuous reduction in resistance; thus, $r^*_{G=1} > r^*_{G=0}$. In turn, lower potentate resistance discontinuously reduces the probability of stopping the ruler's attempt to establish direct rule, $\rho(r^*_{G=1}\gamma T) > \rho(r^*_{G=0}\gamma T)$. The rest of the results attained using the specific functional form for $\rho(\cdot)$ then follow.

### A.3 Marginal Changes in Population

In this section, we examine the how the optimal level of resistance by the local potentate changes with marginal changes to population. Consider the case in which the potentate decides to guard his region against rebellion (i.e., $G = 1$). In this case, the optimal level of resistance decreases with population. This can be seen from condition (A2), which increases as $T$ decreases. Since $\rho(\cdot)$ is concave in $r$ (i.e., $\rho_{rr} < 0$), a higher value of $\rho_r(r^*\gamma T)$ can only be attained when $r^*$ is smaller.

When the potentate guards his region and the cost for the ruler to put down a rebellion is large enough—that is, when $C_R \geq \dfrac{\gamma}{\omega_T(T)}$—then a decline in local population increases the probability that the ruler decides to establish direct rule. This can be seen from condition (A3), which unambiguously increases as population declines. This means that the ruler is willing to establish direct rule for a larger range of values of the investment/administrative cost of deploying direct rule, $C_D$.

We now turn to the case in which the potentate chooses not to guard the region against rebellion (i.e., $G = 0$). In this case, the effect of increasing population becomes ambiguous. Now condition (A2) either increases or decreases its value depending on whether a decrease in population reduces the risk of rebellion more than it reduces the expected tax revenue that the potentate keeps (this can be seen in the denominator of condition (A2)). If the latter dominates, for example, then condition (A2) increases its value, and the optimal resistance by the potentate is less intense as population declines.

The effect of local population changes on the probability that the ruler decides to establish direct rule is also ambiguous, and can be examined in condition (A3). If the local potentate does not guard against rebellion (i.e., $G = 0$), then the ruler does not consider the probability of rebellion when deciding whether to set up direct rule, and expects less tax revenue as population declines. Even if a local drop in population reduces the resistance of the potentate and thus increases the probability of successfully instituting direct rule, capturing all the available tax revenue will not necessarily compensate for the absolute decline in the amount collected. Thus, a marginal decline in population when the potentate chooses not to guard his region has an ambiguous effect on the likelihood of direct rule.

# B. Data Appendix

## B.1 Descriptive Statistics

In this section we present descriptive statistics of the main variables. Table B.1 includes all district-years for the full sample, and the sample with available climate covariates. Table B.2 disaggregates the descriptive statistics for each year in our panel.

Due to the lack of usable tree rings in southeastern Mexico and the Yucatan peninsula, we do not have climate data for the southeastern portion of the sample. These regions account for the gap in observations between groups. Both samples, which correspond to those used in the analysis, are composed of districts with at least two years of data.

**Table B.1:** Descriptive Statistics, All Years

| | Full Sample | | | | | Sample with Climate Covariates | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | count | mean | sd | min | max | count | mean | sd | min | max |
| Direct Rule (%) | 314 | 0.51 | 0.32 | 0 | 1 | 296 | 0.51 | 0.32 | 0 | 1 |
| Population (log) | 314 | 8.59 | 1.34 | 3.99 | 11.7 | 296 | 8.58 | 1.34 | 3.99 | 11.6 |
| Tributaries (log) | 299 | 7.54 | 1.34 | 3.00 | 10.5 | 289 | 7.54 | 1.36 | 3.00 | 10.5 |
| Distance to Mexico City | 314 | 249.6 | 193.6 | 0 | 705.1 | 296 | 231.2 | 182.8 | 0 | 678.7 |
| Avg. elevation | 314 | 1550.5 | 736.3 | 4.97 | 2904.2 | 296 | 1624.5 | 689.3 | 25.8 | 2904.2 |
| Surface area (log) | 314 | 7.65 | 1.25 | 4.68 | 10.9 | 296 | 7.56 | 1.19 | 4.68 | 9.68 |
| Malarial zone | 314 | 0.61 | 0.49 | 0 | 1 | 296 | 0.59 | 0.49 | 0 | 1 |
| Year of European Contact | 314 | 1522.1 | 3.76 | 1518 | 1535 | 296 | 1522.2 | 3.84 | 1518 | 1535 |
| Num. of languages | 309 | 2.80 | 1.64 | 1 | 9 | 294 | 2.80 | 1.68 | 1 | 9 |
| Mine in district by 1600 | 314 | 0.27 | 0.44 | 0 | 1 | 296 | 0.28 | 0.45 | 0 | 1 |
| Drought-rain gap around outbreaks | | | | | | 296 | 0.38 | 0.96 | 0 | 4.22 |
| Drought-rain around outbreaks | | | | | | 296 | 0.14 | 0.35 | 0 | 1 |
| Avg. PDSI | | | | | | 296 | 0.12 | 0.68 | -1.41 | 1.98 |
| Min. PDSI | | | | | | 296 | -3.63 | 0.91 | -5.67 | -1.63 |
| Std. Dev. PDSI | | | | | | 296 | 1.83 | 0.19 | 1.47 | 2.76 |

Estimation sample. We exclude singleton districts; i.e., districts with non-missing data for only one year.

**Table B.2:** Descriptive Statistics, by Year

| | Year: 1550 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Full Sample | | | | | Sample with Climate Covariates | | | | |
| | count | mean | sd | min | max | count | mean | sd | min | max |
| Direct Rule (%) | 51 | 0.42 | 0.36 | 0 | 1 | 45 | 0.43 | 0.36 | 0 | 1 |
| Population (log) | 51 | 9.17 | 1.05 | 6.94 | 11.7 | 45 | 9.13 | 1.03 | 6.94 | 11.6 |
| Tributaries (log) | 42 | 8.14 | 1.06 | 5.92 | 10.5 | 40 | 8.12 | 1.08 | 5.92 | 10.5 |
| Drought-rain gap around outbreaks | | | | | | 45 | 0 | 0 | 0 | 0 |
| Drought-rain around outbreaks | | | | | | 45 | 0 | 0 | 0 | 0 |
| Avg. PDSI | | | | | | 45 | 0.86 | 0.50 | 0.28 | 1.98 |
| Min. PDSI | | | | | | 45 | -2.36 | 0.24 | -2.72 | -1.63 |
| Std. Dev. PDSI | | | | | | 45 | 1.69 | 0.13 | 1.51 | 2.03 |
| | Year: 1570 | | | | | | | | | |
| | Full Sample | | | | | Sample with Climate Covariates | | | | |
| | count | mean | sd | min | max | count | mean | sd | min | max |
| Direct Rule (%) | 112 | 0.46 | 0.32 | 0 | 1 | 107 | 0.46 | 0.32 | 0 | 1 |
| Population (log) | 112 | 9.13 | 1.10 | 5.64 | 11.2 | 107 | 9.15 | 1.12 | 5.64 | 11.2 |
| Tributaries (log) | 110 | 8.12 | 1.11 | 4.62 | 10.2 | 105 | 8.14 | 1.12 | 4.62 | 10.2 |
| Drought-rain gap around outbreaks | | | | | | 107 | 0.52 | 0.96 | 0 | 2.81 |
| Drought-rain around outbreaks | | | | | | 107 | 0.23 | 0.43 | 0 | 1 |
| Avg. PDSI | | | | | | 107 | -0.57 | 0.20 | -1.41 | 0.12 |
| Min. PDSI | | | | | | 107 | -4.36 | 0.38 | -5.07 | -3.16 |
| Std. Dev. PDSI | | | | | | 107 | 1.89 | 0.21 | 1.67 | 2.76 |
| | Year: 1600 | | | | | | | | | |
| | Full Sample | | | | | Sample with Climate Covariates | | | | |
| | count | mean | sd | min | max | count | mean | sd | min | max |
| Direct Rule (%) | 80 | 0.48 | 0.28 | 0 | 1 | 75 | 0.49 | 0.27 | 0 | 1 |
| Population (log) | 80 | 8.46 | 1.26 | 5.42 | 11.4 | 75 | 8.44 | 1.23 | 5.42 | 10.7 |
| Tributaries (log) | 76 | 7.40 | 1.23 | 4.39 | 9.65 | 75 | 7.41 | 1.23 | 4.39 | 9.65 |
| Drought-rain gap around outbreaks | | | | | | 75 | 0.75 | 1.41 | 0 | 4.22 |
| Drought-rain around outbreaks | | | | | | 75 | 0.23 | 0.42 | 0 | 1 |
| Avg. PDSI | | | | | | 75 | 0.080 | 0.39 | -1.04 | 0.57 |
| Min. PDSI | | | | | | 75 | -4.18 | 0.41 | -5.67 | -3.35 |
| Std. Dev. PDSI | | | | | | 75 | 1.80 | 0.15 | 1.47 | 2.25 |
| | Year: 1645 | | | | | | | | | |
| | Full Sample | | | | | Sample with Climate Covariates | | | | |
| | count | mean | sd | min | max | count | mean | sd | min | max |
| Direct Rule (%) | 71 | 0.69 | 0.28 | 0 | 1 | 69 | 0.68 | 0.27 | 0 | 1 |
| Population (log) | 71 | 7.46 | 1.22 | 3.99 | 10.3 | 69 | 7.47 | 1.24 | 3.99 | 10.3 |
| Tributaries (log) | 71 | 6.44 | 1.22 | 3.00 | 9.31 | 69 | 6.44 | 1.23 | 3.00 | 9.31 |
| Drought-rain gap around outbreaks | | | | | | 69 | 0 | 0 | 0 | 0 |
| Drought-rain around outbreaks | | | | | | 69 | 0 | 0 | 0 | 0 |
| Avg. PDSI | | | | | | 69 | 0.77 | 0.25 | 0.52 | 1.53 |
| Min. PDSI | | | | | | 69 | -2.72 | 0.27 | -3.35 | -2.07 |
| Std. Dev. PDSI | | | | | | 69 | 1.88 | 0.17 | 1.67 | 2.48 |

Estimation sample. We exclude singleton districts; i.e., districts with non-missing data for only one year.

## B.2 Balance on Observables in 1550

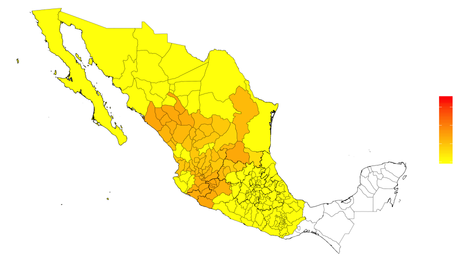
**Table B.3:** Balance on Observables in 1550
Between Districts Affected and Unaffected by Drought-Rain Shocks

| | No Drought-Rain | | Drought-Rain | | Difference | P-value | t-statistic |
|---|---|---|---|---|---|---|---|
| | N | Average | N | Average | | | |
| Direct Rule (%) | 31 | 0.40 | 14 | 0.48 | -0.07 | 0.54 | -0.62 |
| Population (log) | 31 | 9.19 | 14 | 9.00 | 0.19 | 0.58 | 0.56 |
| Tributaries (log) | 31 | 8.16 | 9 | 8.01 | 0.14 | 0.73 | 0.35 |
| Num. of languages | 31 | 2.77 | 14 | 3.00 | -0.23 | 0.70 | -0.40 |
| Any mine | 31 | 0.26 | 14 | 0.29 | -0.03 | 0.85 | -0.19 |
| Malarial zone | 31 | 0.74 | 14 | 0.57 | 0.17 | 0.26 | 1.14 |
| Distance to Mexico City | 31 | 229.19 | 14 | 445.45 | -216.26 | 0.00 | -4.54 |
| Avg. elevation | 31 | 1466.36 | 14 | 1369.10 | 97.25 | 0.63 | 0.491 |
| Surface area (log) | 31 | 7.56 | 14 | 7.45 | 0.11 | 0.79 | 0.27 |

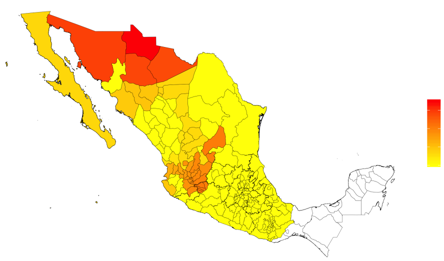
## B.3 Geographic Distribution of Drought-Rain Shocks

**Figure B.1:** Drought-Rain Gap Around Cocoliztli Outbreaks

(a) 1521-1550



(b) 1551-1570



(c) 1571-1600

(d) 1601-1645

# C. Additional Evidence

## C.1 Reduced-Form Results

In this subsection, we present comprehensive econometric evidence on the reduced-form relationship between our climate instruments and the transition to direct rule. The first two columns present estimates using the gap between drought severity and rainfall as the climate measure, and the second two columns do the same with the indicator instrument. In Columns (1) and (3), we report baseline estimates, conditioning on year and district fixed effects only. In Columns (2) and (4) we also include the full set of time-varying and time-interacted controls. In all specifications, the coefficients on the climate variables are positive, indicating that districts experiencing climate conditions conducive to cocoliztli outbreaks saw an increase in the proportion of *encomienda* holdings that transition to direct rule by a given cutoff relative to those that did not. This provides additional evidence in support of the theory and the relevance of the climate instruments.

**Table C.1:** Indigenous Population Collapse and Drought-Rain Around Cocoliztli Outbreaks: Reduced Form

| | Drought-Rain Gap | | Drought-Rain | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Drought-rain around outbreaks | | | 0.085 | $0.16^{**}$ |
| | | | (0.063) | (0.079) |
| Drought-rain gap around outbreaks | $0.036^{*}$ | $0.050^{**}$ | | |
| | (0.021) | (0.022) | | |
| Climate controls | No | Yes | No | Yes |
| Controls $\times$ Year FE | No | Yes | No | Yes |
| Year FE | Yes | Yes | Yes | Yes |
| District FE | Yes | Yes | Yes | Yes |
| Within-District Mean of DV | 0.51 | 0.51 | 0.51 | 0.51 |
| Within-District SD of DV | 0.13 | 0.13 | 0.13 | 0.13 |
| R sq. | 0.82 | 0.83 | 0.82 | 0.83 |
| Observations | 296 | 296 | 296 | 296 |
| Number of districts | 114 | 114 | 114 | 114 |

OLS estimations. The unit-of-analysis is the district-year. Standard errors (clustered at the district level) in parentheses.

## C.2 Empirical Analysis including Singleton Observations

**Table C.2:** Indigenous Population Collapse and Direct Rule: Difference-in-Differences including Singleton Observations

| | Direct Rule (% of District) | | | |
|---|---|---|---|---|
| | Full Sample | Sample with Climate Covariates | | |
| | (1) | (2) | (3) | (4) |
| Population (log) | -0.092** | -0.088* | -0.098* | -0.13*** |
| | (0.044) | (0.045) | (0.050) | (0.043) |
| Climate Controls | No | No | Yes | Yes |
| Controls × Year FE | No | No | Yes | Yes |
| Year of European Contact × Year FE | No | No | No | Yes |
| Initial Population (log) × Year FE | No | No | No | Yes |
| Year FE | Yes | Yes | Yes | Yes |
| District FE | Yes | Yes | Yes | Yes |
| Within-District Mean of DV | 0.50 | 0.51 | 0.51 | 0.51 |
| Within-District SD of DV | 0.14 | 0.13 | 0.13 | 0.13 |
| R sq. | 0.86 | 0.84 | 0.85 | 0.86 |
| Observations | 350 | 319 | 319 | 319 |
| Number of districts | 158 | 137 | 137 | 137 |

OLS estimations. See equation (1) for the econometric specification. The unit-of-analysis is the district-year. Standard errors (clustered at the district level) in parentheses.

**Table C.3:** Tributary Collapse and Direct Rule:
Difference in Differences including Singleton Observations

| | Direct Rule (% of District) | | | |
|---|---|---|---|---|
| | Full Sample | Sample with Climate Covariates | | |
| | (1) | (2) | (3) | (4) |
| Tributaries (log) | -0.083* | -0.081* | -0.10* | -0.12*** |
| | (0.044) | (0.045) | (0.052) | (0.043) |
| Climate Controls | No | No | Yes | Yes |
| Controls × Year FE | No | No | Yes | Yes |
| Year of European Contact × Year FE | No | No | No | Yes |
| Initial Tributaries (log) × Year FE | No | No | No | Yes |
| Year FE | Yes | Yes | Yes | Yes |
| District FE | Yes | Yes | Yes | Yes |
| Within-District Mean of DV | 0.51 | 0.51 | 0.51 | 0.51 |
| Within-District SD of DV | 0.14 | 0.14 | 0.14 | 0.13 |
| R sq. | 0.84 | 0.84 | 0.85 | 0.87 |
| Observations | 321 | 311 | 311 | 301 |
| Number of districts | 140 | 135 | 135 | 131 |

OLS estimations. See equation (1) for the econometric specification. The unit-of-analysis is the district-year. Standard errors (clustered at the district level) in parentheses. In columns (4), the sample shrinks by a couple of observations as a result of missing data on the initial year of European contact in a few districts.‘

## C.3 Empirical Analysis with Tributaries

**Table C.4:** Tributary Collapse and Direct Rule: Difference in Differences

| | Direct Rule (% of District) | | | |
|---|---|---|---|---|
| | Full Sample | Sample with Climate Covariates | | |
| | (1) | (2) | (3) | (4) |
| Tributaries (log) | -0.083* | -0.081* | -0.10** | -0.12*** |
| | (0.043) | (0.043) | (0.050) | (0.041) |
| Climate Controls | No | No | Yes | Yes |
| Controls × Year FE | No | No | Yes | Yes |
| Year of European Contact × Year FE | No | No | No | Yes |
| Initial Tributaries (log) × Year FE | No | No | No | Yes |
| Year FE | Yes | Yes | Yes | Yes |
| District FE | Yes | Yes | Yes | Yes |
| Within-District Mean of DV | 0.51 | 0.51 | 0.51 | 0.50 |
| Within-District SD of DV | 0.14 | 0.14 | 0.14 | 0.13 |
| R sq. | 0.83 | 0.83 | 0.84 | 0.86 |
| Observations | 299 | 289 | 289 | 279 |
| Number of districts | 118 | 113 | 113 | 109 |

OLS estimations. See equation (1) for the econometric specification. The unit-of-analysis is the district-year. Standard errors (clustered at the district level) in parentheses. In columns (4), the sample shrinks by a couple of observations as a result of missing data on the initial year of European contact in a few districts.‘

**Table C.5:** Tributary Collapse and Direct Rule: Instrumental Variables

| | Tributaries (log) | | Direct Rule (% of District) | | Tributaries (log) | | Direct Rule (% of District) | |
|---|---|---|---|---|---|---|---|---|
| | First Stage: OLS | | 2SLS | | First Stage: OLS | | 2SLS | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Tributaries (log) | | | -0.31** | -0.33** | | | -0.30** | -0.30** |
| | | | (0.15) | (0.15) | | | (0.13) | (0.14) |
| Drought-rain gap around outbreaks | -0.13*** | -0.16*** | | | -0.24** | -0.22 | | |
| | (0.041) | (0.060) | | | (0.11) | (0.21) | | |
| Drought-rain around outbreaks | | | | | 0.34 | 0.19 | | |
| | | | | | (0.39) | (0.60) | | |
| Climate controls | No | Yes | No | Yes | No | Yes | No | Yes |
| Controls × Year FE | No | Yes | No | Yes | No | Yes | No | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| District FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Within-District Mean of DV | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 |
| Within-District SD of DV | 0.14 | 0.14 | 0.14 | 0.14 | 0.14 | 0.14 | 0.14 | 0.14 |
| Wald F statistic of excluded instruments | 16.0 | 12.2 | | | 14.0 | 6.29 | | |
| Hansen J statistic | | | . | . | | | 0.0085 | 1.09 |
| Hansen J p-value | | | . | . | | | 0.93 | 0.30 |
| R sq. | 0.93 | 0.95 | 0.12 | 0.21 | 0.93 | 0.95 | 0.13 | 0.25 |
| Observations | 289 | 289 | 289 | 289 | 289 | 289 | 289 | 289 |
| Number of districts | 113 | 113 | 113 | 113 | 113 | 113 | 113 | 113 |

See equations (2) and (3) for the econometric specifications. The unit-of-analysis is the district-year. Standard errors (clustered at the district level) in parentheses.

**Table C.6:** Heterogeneous Effect of Tributary Collapse on Direct Rule, by Rebellion Potential: Difference in Differences

| | Direct Rule (% of District) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Tributaries (log) | -0.066* | -0.095** | -0.11* | -0.14** | -0.12** | -0.16*** | -0.12** | -0.21*** |
| | (0.039) | (0.047) | (0.055) | (0.065) | (0.052) | (0.052) | (0.054) | (0.063) |
| Tributaries (log) × Resistance during conquest | -0.031 | -0.012 | | | | | -0.037 | -0.017 |
| | (0.037) | (0.041) | | | | | (0.037) | (0.038) |
| Tributaries (log) × Num. of languages | | | 0.0090 | 0.014 | | | 0.0061 | 0.016 |
| | | | (0.0091) | (0.011) | | | (0.0078) | (0.010) |
| Tributaries (log) × Num. of towns (1786) | | | | | 0.0040 | 0.0065* | 0.0050 | 0.0085** |
| | | | | | (0.0032) | (0.0036) | (0.0035) | (0.0042) |
| Climate controls | No | Yes | No | Yes | No | Yes | No | Yes |
| Controls × Year FE | No | Yes | No | Yes | No | Yes | No | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| District FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Within-District Mean of DV | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 |
| Within-District SD of DV | 0.14 | 0.14 | 0.14 | 0.13 | 0.14 | 0.14 | 0.14 | 0.13 |
| R sq. | 0.83 | 0.84 | 0.83 | 0.84 | 0.83 | 0.84 | 0.83 | 0.85 |
| Observations | 299 | 289 | 296 | 287 | 299 | 289 | 296 | 287 |
| Number of districts | 118 | 113 | 116 | 112 | 118 | 113 | 116 | 112 |

OLS estimations. See equation (A1) for the econometric specification. The unit-of-analysis is the district-year. Standard errors (clustered at the district level) in parentheses.

**Table C.7:** Heterogeneous Effect of Tributary Collapse on Direct Rule, by Outside Encomendero Options: Difference in Differences

| | Direct Rule (% of District) | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Tributaries (log) | -0.064* | -0.083* | | |
| | (0.037) | (0.046) | | |
| Tributaries (log) × Any mine | -0.060 | -0.070 | | |
| | (0.052) | (0.056) | | |
| Any mine × 1570 | | | 0.18 | 0.22 |
| | | | (0.12) | (0.13) |
| Any mine × 1600 | | | 0.20 | 0.23 |
| | | | (0.13) | (0.14) |
| Any mine × 1645 | | | 0.23* | 0.26* |
| | | | (0.14) | (0.14) |
| Climate controls | No | Yes | No | Yes |
| Controls × Year FE | No | Yes | No | Yes |
| Year FE | Yes | Yes | Yes | Yes |
| District FE | Yes | Yes | Yes | Yes |
| Within-District Mean of DV | 0.51 | 0.51 | 0.51 | 0.51 |
| Within-District SD of DV | 0.14 | 0.14 | 0.14 | 0.13 |
| R sq. | 0.83 | 0.85 | 0.83 | 0.84 |
| Observations | 299 | 289 | 314 | 296 |
| Number of districts | 118 | 113 | 122 | 114 |

OLS estimations. See equation (A1) for the econometric specification. The unit-of-analysis is the district-year. Standard errors (clustered at the district level) in parentheses.

## C.4 Empirical Analysis with 10-Year Window for Cutoffs

**Table C.8:** Indigenous Population Collapse and Direct Rule: Difference in Differences

| | Direct Rule (% of District) | | | |
|---|---|---|---|---|
| | Full Sample | Sample with Climate Covariates | | |
| | (1) | (2) | (3) | (4) |
| Population (log) | -0.12*** | -0.089** | -0.096** | -0.12*** |
| | (0.038) | (0.044) | (0.047) | (0.040) |
| Climate Controls | No | No | Yes | Yes |
| Controls × Year FE | No | No | Yes | Yes |
| Year of European Contact × Year FE | No | No | No | Yes |
| Initial Population (log) × Year FE | No | No | No | Yes |
| Year FE | Yes | Yes | Yes | Yes |
| District FE | Yes | Yes | Yes | Yes |
| Within-District Mean of DV | 0.47 | 0.51 | 0.51 | 0.51 |
| Within-District SD of DV | 0.13 | 0.13 | 0.13 | 0.13 |
| R sq. | 0.85 | 0.83 | 0.84 | 0.85 |
| Observations | 380 | 307 | 307 | 307 |
| Number of districts | 138 | 117 | 117 | 117 |

OLS estimations. See equation (1) for the econometric specification. The unit-of-analysis is the district-year. Standard errors (clustered at the district level) in parentheses.

**Table C.9:** Indigenous Population Collapse and Direct Rule: Difference in Differences with Tributary Measure

| | Direct Rule (% of District) | | | |
|---|---|---|---|---|
| | Full Sample | Sample with Climate Covariates | | |
| | (1) | (2) | (3) | (4) |
| Tributaries (log) | -0.086* | -0.084* | -0.10** | -0.12*** |
| | (0.044) | (0.044) | (0.050) | (0.040) |
| Climate Controls | No | No | Yes | Yes |
| Controls × Year FE | No | No | Yes | Yes |
| Year of European Contact × Year FE | No | No | No | Yes |
| Initial Tributaries (log) × Year FE | No | No | No | Yes |
| Year FE | Yes | Yes | Yes | Yes |
| District FE | Yes | Yes | Yes | Yes |
| Within-District Mean of DV | 0.51 | 0.51 | 0.51 | 0.50 |
| Within-District SD of DV | 0.14 | 0.13 | 0.13 | 0.13 |
| R sq. | 0.83 | 0.83 | 0.84 | 0.86 |
| Observations | 309 | 299 | 299 | 289 |
| Number of districts | 121 | 116 | 116 | 112 |

OLS estimations. See equation (1) for the econometric specification. The unit-of-analysis is the district-year. Standard errors (clustered at the district level) in parentheses.

**Table C.10:** Indigenous Population Collapse and Direct Rule: Instrumental Variables

| | Population (log) | | Direct Rule (% of District) | | Population (log) | | Direct Rule (% of District) | |
|---|---|---|---|---|---|---|---|---|
| | First Stage: OLS | | 2SLS | | First Stage: OLS | | 2SLS | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Population (log) | | | -0.25** | -0.32** | | | -0.25** | -0.30** |
| | | | (0.12) | (0.14) | | | (0.12) | (0.14) |
| Drought-rain gap around outbreaks | -0.14*** | -0.15** | | | -0.19* | -0.19 | | |
| | (0.041) | (0.063) | | | (0.11) | (0.19) | | |
| Drought-rain around outbreaks | | | | | 0.14 | 0.13 | | |
| | | | | | (0.37) | (0.54) | | |
| Climate controls | No | Yes | No | Yes | No | Yes | No | Yes |
| Controls × Year FE | No | Yes | No | Yes | No | Yes | No | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| District FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Within-District Mean of DV | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 |
| Within-District SD of DV | 0.13 | 0.13 | 0.13 | 0.13 | 0.13 | 0.13 | 0.13 | 0.13 |
| Wald F statistic of excluded instruments | 19.3 | 9.93 | | | 14.0 | 5.05 | | |
| Hansen J statistic | | | . | . | | | 0.071 | 1.51 |
| Hansen J p-value | | | . | . | | | 0.79 | 0.22 |
| R sq. | 0.93 | 0.94 | 0.23 | 0.19 | 0.93 | 0.94 | 0.22 | 0.23 |
| Observations | 307 | 307 | 307 | 307 | 307 | 307 | 307 | 307 |
| Number of districts | 117 | 117 | 117 | 117 | 117 | 117 | 117 | 117 |

See equations (2) and (3) for the econometric specifications. The unit-of-analysis is the district-year. Standard errors (clustered at the district level) in parentheses.

**Table C.11:** Tributary Collapse and Direct Rule: Instrumental Variables with Tributary Measure

| | Tributaries (log) | | Direct Rule (% of District) | | Tributaries (log) | | Direct Rule (% of District) | |
|---|---|---|---|---|---|---|---|---|
| | First Stage: OLS | | 2SLS | | First Stage: OLS | | 2SLS | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Tributaries (log) | | | -0.33** | -0.37** | | | -0.32** | -0.34** |
| | | | (0.16) | (0.17) | | | (0.14) | (0.16) |
| Drought-rain gap around outbreaks | -0.12*** | -0.15** | | | -0.23** | -0.20 | | |
| | (0.040) | (0.063) | | | (0.11) | (0.21) | | |
| Drought-rain around outbreaks | | | | | 0.33 | 0.16 | | |
| | | | | | (0.39) | (0.59) | | |
| Climate controls | No | Yes | No | Yes | No | Yes | No | Yes |
| Controls $\times$ Year FE | No | Yes | No | Yes | No | Yes | No | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| District FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Within-District Mean of DV | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 |
| Within-District SD of DV | 0.13 | 0.13 | 0.13 | 0.13 | 0.13 | 0.13 | 0.13 | 0.13 |
| Wald F statistic of excluded instruments | 15.4 | 9.31 | | | 13.2 | 4.72 | | |
| Hansen J statistic | | | . | . | | | 0.017 | 0.96 |
| Hansen J p-value | | | . | . | | | 0.90 | 0.33 |
| R sq. | 0.94 | 0.95 | 0.055 | 0.12 | 0.94 | 0.95 | 0.074 | 0.17 |
| Observations | 299 | 299 | 299 | 299 | 299 | 299 | 299 | 299 |
| Number of districts | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 |

See equations (2) and (3) for the econometric specifications. The unit-of-analysis is the district-year. Standard errors (clustered at the district level) in parentheses.

## C.5 Empirical Analysis with Balanced Panel

**Table C.12:** Indigenous Population and Tributary Collapse and Direct Rule: Difference in Differences with Balanced Panel

| | Direct Rule (% of District) | | | | | |
|---|---|---|---|---|---|---|
| | Balanced Panel Population (log) | | | Balanced Panel Tributaries (log) | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Population (log) | -0.21** | -0.14 | -0.29*** | | | |
| | (0.093) | (0.14) | (0.084) | | | |
| Tributaries (log) | | | | -0.23 | -0.24 | -0.48*** |
| | | | | (0.13) | (0.17) | (0.12) |
| Climate Controls | No | Yes | Yes | No | Yes | Yes |
| Controls $\times$ Year FE | No | Yes | Yes | No | Yes | Yes |
| Year of European Contact $\times$ Year FE | No | No | Yes | No | No | Yes |
| Initial Population (log) $\times$ Year FE | No | No | Yes | | | |
| Initial Tributaries (log) $\times$ Year FE | | | | No | No | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| District FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Within-District Mean of DV | 0.43 | 0.43 | 0.43 | 0.39 | 0.39 | 0.39 |
| Within-District SD of DV | 0.17 | 0.17 | 0.17 | 0.17 | 0.17 | 0.17 |
| R sq. | 0.86 | 0.91 | 0.98 | 0.84 | 0.90 | 0.99 |
| Observations | 48 | 48 | 48 | 40 | 40 | 40 |
| Number of districts | 12 | 12 | 12 | 10 | 10 | 10 |

OLS estimations. See equation (1) for the econometric specification. The unit-of-analysis is the district-year. Standard errors (clustered at the district level) in parentheses.

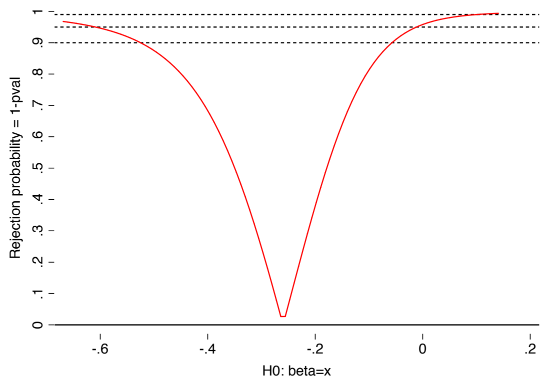
## C.6 Weak-Instrument Robust Inference

Though first-stage Wald F-statistics are above typical thresholds for concern over weak instruments, these rules of thumb, as well as the standard cutoffs developed in Stock and Yogo (2005), are based on independent, identically distributed errors and may not be appropriate when errors have a clustered structure. Clustering may compound bias due to weak instruments (Cameron and Miller, 2015).[4] To address this possibility, we also estimate our models using the Anderson-Rubin (AR) method, which is robust to weak instruments and can be generalized for cluster-robust inference (Cameron and Miller, 2015). The $(1-\alpha)\%$ AR confidence interval is constructed by inverting the AR weak instrument test of size $\alpha$ and identifying the values of $\beta^*$ for which the joint null of $\beta = \beta^*$ and $E(Zu) = 0$ cannot be rejected (e.g., Stock and Yogo, 2005). In all four of our IV specifications (one and two instruments with and without the vector of controls), the 90% AR confidence intervals contain strictly negative values. In the 95% case, this is true in all but one specification. This adds to our confidence that the coefficient on population is negative, as suggested by the theory.

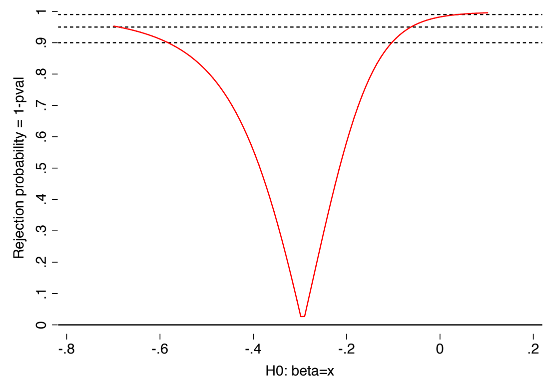
Graphs of the 90%, 95%, and 99% AR confidence intervals are presented in Figure C.1. Plotted are the rejection probabilities of the joint null described above. The dotted lines represent the appropriate cutoffs for the three confidence levels. The AR confidence interval is the region where this line lies below the appropriate cutoffs.

[4] Asymptotics are in the number of clusters rather than the number of observations in this case (Cameron and Miller, 2015).

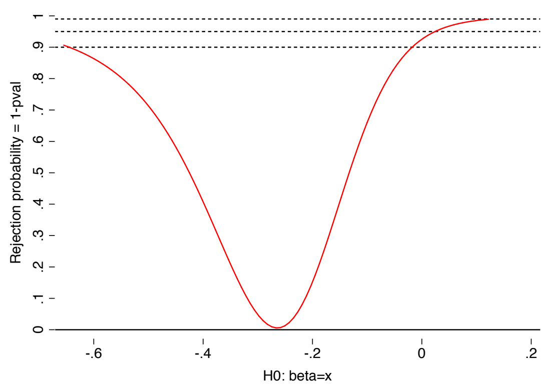
**Figure C.1:** Anderson-Rubin Test Confidence Intervals



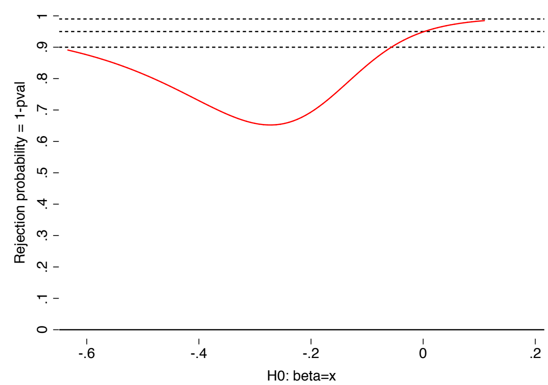

(a) Single instrument, no controls




(b) Single instrument, controls




(c) Two instruments, no controls




(d) Two instruments, controls

## C.7 Alternative Instrument Construction

The results in Table 2 rely on two instruments: first, an indicator for whether a given district experienced a severe, longer than two-year drought that ended 1-2 years prior to any outbreak of cocoliztli in the period; second, the magnitude of the rain-drought gap, i.e., the numeric difference between the lowest PDSI recorded in a drought ending 1-2 years prior to a cocoliztli outbreak and the PDSI of the first non-drought year.

**Table C.13:** Indigenous Population Collapse and Direct Rule: Instrumental Variables with Alternative Instruments

| | Population (log) | | Direct Rule (% of District) | | Population (log) | | Direct Rule (% of District) | |
|---|---|---|---|---|---|---|---|---|
| | First Stage: OLS | | 2SLS | | First Stage: OLS | | 2SLS | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Population (log) | | | -0.33** | -0.42** | | | -0.23* | -0.12 |
| | | | (0.16) | (0.17) | | | (0.12) | (0.11) |
| Drought-rain gap | -0.12** | -0.14** | | | -0.29** | -0.36** | | |
| | (0.045) | (0.055) | | | (0.12) | (0.16) | | |
| Any Drought-rain | | | | | 0.52 | 0.66 | | |
| | | | | | (0.41) | (0.40) | | |
| Climate controls | No | Yes | No | Yes | No | Yes | No | Yes |
| Controls × Year FE | No | Yes | No | Yes | No | Yes | No | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| District FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Within-District Mean of DV | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 |
| Within-District SD of DV | 0.13 | 0.13 | 0.13 | 0.13 | 0.13 | 0.13 | 0.13 | 0.13 |
| Wald F statistic of excluded instruments | 10.9 | 10.9 | | | 15.5 | 6.54 | | |
| Hansen J statistic | | | . | . | | | 1.00 | 7.70 |
| Hansen J p-value | | | . | . | | | 0.32 | 0.0055 |
| R sq. | 0.93 | 0.94 | 0.071 | 0.0029 | 0.93 | 0.94 | 0.27 | 0.40 |
| Observations | 296 | 296 | 296 | 296 | 296 | 296 | 296 | 296 |
| Number of districts | 114 | 114 | 114 | 114 | 114 | 114 | 114 | 114 |

See equations (2) and (3) for the econometric specifications. The unit-of-analysis is the district-year. Standard errors (clustered at the district level) in parentheses.

In this appendix, we assess the robustness of the results to an alternative specification of the instruments that does not rely on information about the timing of cocoliztli outbreaks. Specifically, these new instruments use information on all droughts during the early colonial period, and not just those that ended just prior to a cocoliztli outbreak. The instruments are constructed similarly to those in the main text: an indicator for whether a district experienced a severe (i.e., longer than 2 years) droughts and the largest numeric difference between the lowest PDSI during any drought and the PDSI of the first non-drought year. Severe droughts are relatively rare in the data outside of cocoliztli years. These new variables identify only four new drought episodes that were not captured in the main IV specifications.

The estimates using this new set of instruments are presented in Table C.13. The first-stage results generally provide evidence of instrument relevance. The first-stage Wald statistics are above standard rules-of-thumb when using only the drought-rain gap instrument. When using both instruments, the first stage relationship is weaker after including the full set of controls (column 6). The IV coefficient estimates are negative and are of comparable magnitude to those of our main results in all but one specification (column 8), where the magnitude is roughly halved and the coefficient is not statistically significant.[5] Overall, however, the results are nearly identical to those using the main instruments, providing further confidence in the relationship between population collapse and the transition to direct rule suggested by the theory.

[5]Hansen J-statistics of overidentification restrictions are smaller than in the main specifications, especially that of column 8. We believe that it is likely that the two instruments are picking up different components of the climate shock treatment: the exposure to a drought-rain sequence itself and the swing to excess rainfall conducive to the spread of rodent-transmitted pathogens respectively. As many have noted, J-statistics will be large under these conditions. They cannot and do not provide direct evidence on the exclusion restrictions, which are untestable (e.g., Deaton, 2010; Parente and Silva, 2012).

## D. Additional Evidence: Mechanisms

In this section, we provide additional empirical support for our theory. First, we provide econometric evidence on the auxiliary hypotheses, discussed in Sections 1 and 4.5. To do this, we adapt our difference-in-differences estimation strategy to examine heterogeneity in the effect of the population collapse. Specifically, we amend equation 1 to estimate:

$$DirectRule_{it} = \beta_1 lnPop_{it} + \beta_2 lnPop_{it} M_i + \Theta_t X_i + \Pi U_{it} + \lambda_t + \gamma_i + \varepsilon_{it}, \qquad \text{(A1)}$$

where $M_i$ is a district-level measure of either the population's ability to coordinate rebellion (for Hypothesis 2), or the value of *encomenderos'* outside earning options (for Hypotheses 3 and 4). Because we do not have a suitable instrument for $M_i$, we cannot adopt the instrumental-variables strategy outlined in the prior section. However, the difference-in-differences approach still enables us to provide suggestive evidence on heterogeneity, controlling for time-invariant and time-varying district characteristics and common temporal trends across districts. We therefore include the same vector of time-varying and time-interacted controls as in the baseline models.

We measure the potential to organize a rebellion using three variables: whether a district violently resisted Spanish conquest at first contact, the number of indigenous languages spoken the district, and the number of towns present in that district in 1786.[6] These measures were digitized from Gerhard (1993*a*,*b*,*c*). We argue that threat of rebellion should be higher in districts with a history of resistance to Spanish control, as well as in places where coordination is not complicated by the existence of numerous unintelligible languages or a large number of dispersed settlements. This coordination argument is supported by historical work on our context (e.g., Katz, 1988; Gerhard, 1993*a*), and by a broader literature on homogeneity, population density, and collective action (e.g., Homer-Dixon, 1999; Salehyan and Gleditsch, 2006). From hypothesis 2, we therefore expect that $\beta_2 > 0$ when $M_i$ measures the number of languages or towns.

[6]We note that, if interpreted causally, the estimates of equation A1 could be subject to post-treatment bias when including the number of towns in 1786 given that our measure was recorded after the population collapse. There is a strong correlation in the overall concentration of population across districts before the collapse and following

**Table D.1:** Heterogeneous Effect of Indigenous Population Collapse on Direct Rule, by Rebellion Potential: Difference in Differences

| | Direct Rule (% of District) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Population (log) | -0.072* | -0.085* | -0.12** | -0.14** | -0.13*** | -0.16*** | -0.13** | -0.21*** |
| | (0.038) | (0.046) | (0.055) | (0.061) | (0.048) | (0.052) | (0.051) | (0.061) |
| Population (log) × Resistance during conquest | -0.033 | -0.021 | | | | | -0.038 | -0.025 |
| | (0.036) | (0.040) | | | | | (0.034) | (0.036) |
| Population (log) × Num. of languages | | | 0.0095 | 0.014 | | | 0.0064 | 0.016 |
| | | | (0.0092) | (0.011) | | | (0.0080) | (0.0099) |
| Population (log) × Num. of towns (1786) | | | | | 0.0048 | 0.0072* | 0.0056* | 0.0093** |
| | | | | | (0.0032) | (0.0038) | (0.0033) | (0.0042) |
| Climate controls | No | Yes | No | Yes | No | Yes | No | Yes |
| Controls × Year FE | No | Yes | No | Yes | No | Yes | No | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| District FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Within-District Mean of DV | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 |
| Within-District SD of DV | 0.14 | 0.13 | 0.14 | 0.13 | 0.14 | 0.13 | 0.14 | 0.13 |
| R sq. | 0.84 | 0.84 | 0.84 | 0.84 | 0.84 | 0.84 | 0.84 | 0.85 |
| Observations | 314 | 296 | 309 | 294 | 314 | 296 | 309 | 294 |
| Number of districts | 122 | 114 | 120 | 113 | 122 | 114 | 120 | 113 |

OLS estimations. See equation (A1) for the econometric specification. The unit-of-analysis is the district-year. Standard errors (clustered at the district level) in parentheses.

Our results evaluating Hypothesis 2 are presented in Table D.1. In the first column, we use an indicator for resistance to conquest at first contact. We present estimates of the heterogeneous effect of the population collapse on the adoption of direct rule, including district and year fixed effects. In the second column, we repeat the analysis including the full set of control variables. Columns (2) through (6) estimate analogous models using the number of languages spoken in the district and the number of settlements in 1786 as alternative proxies for the difficulty of collective action. In columns (7) and (8), we use all three measures of rebellion potential. Across models, districts with higher rebellion potential, including those with a history of resistance as well as with fewer obstacles to coordination—those where fewer languages are spoken and where the population is distributed into fewer distinct towns—display a magnified effect of a decline in population on direct rule adoption. The estimates of heterogeneous effect are consistent in magnitude across models, though they are only precisely estimated for the number of towns in models (6)-(8) and for the number of languages in model (8).

To assess Hypotheses 3 and 4 on heterogeneity by the availability of outside earnings options for *encomenderos*, we digitized and geocoded information on the placement of mines in 1600 from UNAM (2007). Mining was arguably the engine of New Spain's economy and represented one of the major sources of income for elites (Knight, 2002). We code whether a district contains a mine as a measure of possible opportunities for wealth extraction from this sector. Our theory holds that an *encomendero* should be less likely to invest in defending his district from rebellion or resisting the Crown's attempts to centralize power as the value of his outside option increases. Because of this, the Crown should be more willing and able to establish direct rule in areas with mines. This suggests that $\beta_2 < 0$ when $M_i$ measures mining presence in the district. To evaluate Hypothesis 4, which states that more valuable outside options should increase the likelihood of direct rule, we simply interact the mine indicator with each year indicator to examine the relative level of direct rule

the recovery of Mexico's indigenous population in the 17th century (Sellars and Alix-Garcia, 2018). However, this subsection is intended to provide suggestive evidence on the theory, and the results should be interpreted with caution.

adoption in mining and non-mining districts in each year. Our theory predicts that the coefficient on that interaction term should be positive for all years.

**Table D.2:** Heterogeneous Effect of Indigenous Population Collapse on Direct Rule, by Outside Encomendero Options: Difference in Differences

| | Direct Rule (% of District) | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Population (log) | -0.066* | -0.074* | | |
| | (0.036) | (0.044) | | |
| Population (log) × Any mine | -0.067 | -0.074 | | |
| | (0.047) | (0.051) | | |
| Any mine × 1570 | | | 0.18 | 0.22 |
| | | | (0.12) | (0.13) |
| Any mine × 1600 | | | 0.20 | 0.23 |
| | | | (0.13) | (0.14) |
| Any mine × 1645 | | | 0.23* | 0.26* |
| | | | (0.14) | (0.14) |
| Climate controls | No | Yes | No | Yes |
| Controls × Year FE | No | Yes | No | Yes |
| Year FE | Yes | Yes | Yes | Yes |
| District FE | Yes | Yes | Yes | Yes |
| Within-District Mean of DV | 0.51 | 0.51 | 0.51 | 0.51 |
| Within-District SD of DV | 0.14 | 0.13 | 0.14 | 0.13 |
| R sq. | 0.84 | 0.85 | 0.83 | 0.84 |
| Observations | 314 | 296 | 314 | 296 |
| Number of districts | 122 | 114 | 122 | 114 |

OLS estimations. See equation (A1) for the econometric specification. The unit-of-analysis is the district-year. Standard errors (clustered at the district level) in parentheses.

The first two columns in Table D.2 present suggestive evidence on Hypothesis 3, which holds that the effect of a decline in population on the transition to direct rule should be larger in the presence of better outside options. As predicted, our estimates suggest that the effect of a decline in population was greater in mining areas, though this heterogeneous effect is not precisely estimated. Columns (3) and (4) provide supportive evidence on Hypothesis 4, which addresses the overall level of direct rule adoption. As expected, districts with mines are more likely to adopt direct rule in every period, conditional on covariates.

As a final way of building support for our theory, we consider alternative theories about the design of the *encomienda* and the transition to direct rule in colonial Mexico. In a notable debate, Yeager (1995) and Pastore (1998) present competing theories about the institutional design of the encomienda in the Americas. As in our argument, Yeager places a central emphasis on the Crown's security concerns in explaining both the design of the *encomienda* and the transition from *encomienda* to *corregimiento*. His emphasis is, however, somewhat narrower than ours. He argues that the major threat to Crown security came primarily from elites—the *encomenderos*—and that royal officials therefore chose to absorb the holdings of the most powerful *encomenderos* first to undermine their threat to rule.[7] Because tribute wealth was directly linked to population, he predicts that the Crown should begin with the most populous holdings in the transition to direct rule. Using data from Gibson (1964), Yeager shows that this pattern held in the area around Mexico City: the largest *encomiendas* in population were brought under direct rule first. We replicate Yeager's findings in Table D.3.

**Table D.3:** Additional Evidence: Replication of Yeager (1995)

| | Year of Direct Rule<br>Full Sample<br>(1) | Year of Direct Rule<br>Non-Perpetual<br>(2) |
|---|---|---|
| trib_1560 | -0.0095$^{***}$ | -0.0087$^{***}$ |
| | (0.0013) | (0.0012) |
| Constant | 1656.5$^{***}$ | 1638.3$^{***}$ |
| | (14.1) | (10.6) |
| R sq. | 0.24 | 0.39 |
| Observations | 36 | 33 |

OLS estimations. The unit of analysis is the *encomienda*. Huber-White robust standard errors in parentheses. Data from Gibson (1964) and Yeager (1995).

Our work shows that the transition to direct rule happened faster where the population declined

[7] In our theory, we emphasize *encomiendero* resistance but also the threat of rebellion from below.

more precipitously, whereas Yeager argues that more populous holdings were brought under direct rule by the Crown first. Given the different role of population in our theories and these potentially contradictory empirical findings, it is worth discussing how Yeager's work relates to ours. Several features of Yeager's analysis differ from ours. First, Yeager exclusively examines holdings in the central area of New Spain near what became Mexico City. As discussed in Section 2, Mexico City was the heart of Spanish power in the Americas, and this region was among the first to be brought under solid Spanish control. The costs of pacification were therefore much lower in this region relative to others, obviating the need to rely on local elites to secure political control. Yeager's analysis therefore examines an area where the attractiveness of indirect rule would be especially low under our theory: there are few gains from outsourcing the costs of providing security to elites, while the potential benefits to absorbing the holdings are high. We condition on both distance to Mexico City and district by year in our analysis in part to address this concern. However, the broader regional trends in the transition to direct rule are poorly explained by Yeager's elite-centered argument. The transition to direct rule was very slow in regions like Nuevo Leon and the Yucatan where elite power and extraction were especially high. We argue that the high threat of rebellion from below in these regions can explain why the Crown continued to rely on indirect rule in these regions. As Pastore (1998) notes, while the *encomenderos* could pose a security threat to the Crown, royal officials seemed more concerned about the threat of generalized rebellion from below or security threats from other empires. This explains why indirect rule lasted the longest in strategically important and difficult to subdue frontier areas throughout the Empire (p. 513–4).

An additional difference between our work and that of Yeager is the unit of analysis. Yeager's analysis was conducted at the level of the holding or *encomienda*, whereas ours is conducted at the level of the district. It would be impossible to conduct our analysis at a lower level of aggregation given that panel data on population are unavailable below the district level except in very few areas. We also believe that the district is the correct unit of analysis to assess our theory, which hinges on the changing threat of rebellion from below in areas that lose population. Each district typically contained numerous *encomiendas*—sometimes well over a dozen—of differing sizes. Most holdings mapped onto a single village or a handful of villages, and some cut across villages to specific neighborhoods or sections of settlements, especially in the Yucatan. Any uprising large enough to worry the Crown would therefore have to cut across several individual holdings. For these reasons, we believe that the threat of rebellion is best assessed at the district level.

Nonetheless, Yeager's finding that the largest *encomiendas* within a region were taken under direct rule first does not itself contradict our theory. Assuming, as we do, that the central ruler seeks to maximize revenue for the Crown, he should choose to seize the most profitable (and therefore the most populous) holdings within a given district once political control has been established. This is broadly consistent with the pattern seen in Mexico: in areas that had been brought under solid Crown control, the largest *encomiendas* seem to have transition into direct rule first (Gibson, 1964; Zavala, 1973; Garcia Bernal, 1979). Taken together, our findings at the district level and Yeager's at the *encomienda* level help to rule out common alternative explanations for the pattern of transition to direct rule. Pastore (1998), for example, notes that Yeager's theory is observationally equivalent with a simple revenue maximization story given that both predict that populous indirect-rule holdings would be absorbed by the Crown first. However, both our findings and broader regional trends in Mexico are not consistent with this simple account of revenue maximization. Conversely, Yeager's finding that the Crown targeted larger *encomiendas* within districts, confirmed qualitatively by Gibson and Zavala, is inconsistent with the story that the Crown targeted "underperforming" districts for absorption. Our theory emphasizing the tradeoff between maximizing revenue collection and maintaining political control, however, is consistent with both of these findings.

A few other alternative explanations for our findings are worth considering as well. First, a potential concern is that areas that had been brought under Crown control may have disproportionately suffered from epidemics, inducing a relationship between the population collapse and the transition to direct rule. It is not the case that areas under solid control experienced the worst declines in population in these epidemics. For example, the worst-hit areas for the severe 1570–6 cocoliztli epidemic were in sections of what are now Durango and far southern Michoacan and Jalisco states, which were frontier areas at this time (Acuña Soto, Calderon Romero and Maguire, 2000; Acuña Soto et al., 2002). In addition, because we include district-by-year controls in our empirical models, we are examining differential trends toward direct rule within regions.

Another alternative explanation could be that the Crown was more likely to transition to direct rule in disease-affected regions because of concerns about the exploitation of the indigenous population in population-scarce areas. While it is true that some royal officials and clergy were concerned with protecting the indigenous population (famously Fr. Bartolomé de las Casas), it is unclear whether *corregimientos* improved the treatment of the indigenous population in practice (Gibson, 1964). Furthermore, the overall pattern of the transition to direct rule is not consistent with this alternative explanation. Some areas that were both population-scarce and hit hard by disease, especially parts of the far northeast, retained a particularly violent form of the *encomienda* for centuries because of the persistent threat of rebellion and the lack of elite outside options in this region (e.g., Gerhard, 1993*b*). While the Crown could have had incentives to protect indigenous communities in certain situations, these are not necessarily related to population scarcity (Franco-Vivanco, 2017).

A final concern is that areas suffering from epidemics during the 1550–1645 period may have had a different prior experience with disease, potentially contaminating our analysis. Given uncertainty about the characteristics of cocoliztli, it is not clear whether prior exposure should lead to increased later mortality (because certain areas are more susceptible to disease) or decreased later mortality (because susceptible populations had been affected in the earlier epidemics). However, the instrumental-variables empirical strategy we adopt leverages differences in the climatic conditions associated with cocoliztli rather than data on the epidemics themselves. This, along with our fixed-effects estimation strategy helps to avoid potential confounds with mortality before the 1550 period.

# E. Additional Evidence: Persistent Effects of Early Direct Rule

**Table E.1:** Present-Day State Presence and Exposure to Early Direct Rule

| | Number of Postal Offices | | Number of Offices of the Public Prosecutor | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Years under direct rule (1521-1645) | 0.0022$^{**}$ | 0.0026$^{***}$ | 0.010 | 0.012$^{*}$ |
| | (0.0011) | (0.00097) | (0.0073) | (0.0072) |
| Population in 2015 (log) | | 0.54$^{***}$ | | 2.38$^{***}$ |
| | | (0.077) | | (0.63) |
| Surface area (log) | | -0.099$^{**}$ | | -0.78$^{**}$ |
| | | (0.049) | | (0.38) |
| Geographic controls | Yes | Yes | Yes | Yes |
| State indicators | Yes | Yes | Yes | Yes |
| Mean of DV | 0.59 | 0.59 | 1.16 | 1.16 |
| SD of DV | 1.31 | 1.31 | 8.87 | 8.87 |
| Mean of years under direct rule | 64.0 | 64.0 | 64.0 | 64.0 |
| SD of years under direct rule | 31.1 | 31.1 | 31.1 | 31.1 |
| R sq. | 0.28 | 0.45 | 0.048 | 0.12 |
| Number of Municipios | 1908 | 1907 | 1908 | 1907 |

OLS estimations. The unit-of-analysis is the municipio. Huber-White robust standard errors standard errors in parentheses. The number of years under direct rule is weighted the proportion of *encomiendas* brought under direct rule in each cutoff. We weight by surface area when aggregating to the municipio level. Data from postal offices from Correos de México (2017); data from the Public Prosecutor's offices and agents from INEGI (2016).

## F. Appendix References


Acuña Soto, Rodolfo, David W. Stahle, Malcolm K. Cleaveland and Matthew D. Therrell. 2002. "Megadrought and Megadeath in 16th Century Mexico." *Emerging Infectious Diseases* 8(4):360–362.

Acuña Soto, Rodolfo, Leticia Calderon Romero and James H. Maguire. 2000. "Large Epidemics of Hemorrhagic Fevers in Mexico, 1545-1815." *American Journal of Tropical Medicine and Hygiene* 62(6):733–739.

Boserup, Ester. 1965. *The Conditions of Agricultural Growth: The Economics of Agrarian Change under Population Pressure*. Oxford.

Cameron, A. Colin and Douglas L. Miller. 2015. "A Practioner's Guide to Cluster-Robust Inference." *Journal of Human Resources* 50(2):317–373.

Correos de México. 2017. "Cobertura por Oficina.".

Deaton, Angus. 2010. "Instruments, Randomization, and Learning about Development." *Journal of Economic Literature* 48:424–455.

Franco-Vivanco, Edgar. 2017. "Redistributive Justice and Social Order: Indigenous Claims in the Courts of Colonial Mexico." *Working Paper* .

Garcia Bernal, Manuela Cristina. 1979. *Población y encomienda en Yucatán bajo los austrias*. Sevilla, ES: Escuela de Estudios Hispano-Americanos.

Gerhard, Peter. 1993*a*. *A guide to the Historical Geography of New Spain: Revised Edition*. University of Oklahoma Press.

Gerhard, Peter. 1993*b*. *The North Frontier of New Spain*. Norman, OK: University of Oklahoma Press.

Gerhard, Peter. 1993*c*. *The Southeast Frontier of New Spain*. Norman, OK: University of Oklahoma Press.

Gibson, Charles. 1964. *The Aztecs Under Spanish Rule: A History of the Indians of the Valley of Mexico, 1519-1810*. Palo Alto, CA: Stanford University Press.

Goldstone, Jack A. 2002. "Population and Security: How Demographic Change Can Lead to Violent Conflict." *Journal of International Affairs* 56(1):3–21.

Homer-Dixon, Thomas. 1999. *Environment, Scarcity, and Violence*. Princeton, NJ: Princeton University Press.

INEGI. 2016. "Censo Nacional de Procuración de Justicia Estatal 2016.".

Katz, Friedrich. 1988. Rural Uprisings in Preconquest and Colonial Mexico. In *Riot, Rebellion, and Revolution: Rural Social Conflict in Mexico*. Princeton, NJ: Princeton University Press.

Knight, Alan. 2002. *Mexico: The Colonial Era*. New York, NY: Cambridge University Press.

Parente, Paulo M.D.C. and J.M.C. Santos Silva. 2012. "A cautionary note on tests of overidentifying restrictions." *Economics Letters* 115:314–317.

Pastore, Mario. 1998. "Government, Taxation, Coercion, and Ideology: A Comment on Yeager." *The Journal of Economic History* 58(2):511–520.

Salehyan, Idean and Kristian Skrede Gleditsch. 2006. "Refugees and the Spread of War." *International Organization* 60:335–366.

Sellars, Emily A. and Jennifer Alix-Garcia. 2018. "Labor Scarcity, Land Tenure, and Historical Legacy: Evidence from Mexico." *Journal of Development Economics* 135:504–516.

Stock, James H. and Motohiro Yogo. 2005. Testing for Weak Instruments in Linear IV Regression. In *Identification and Inference for Econometric Models: Essays in Honor of Thomas J. Rothenberg*, ed. Donald W.K. Andrews and James H. Stock. New York: Cambridge University Press.

Taylor, William B. 1979. *Drinking, Homicide, and Rebellion in Colonial Mexican Villages*. Stanford, CA: Stanford University Press.

UNAM. 2007. *Nuevo atlas nacional de México*. México, DF: UNAM.

Yeager, Timothy J. 1995. "Encomienda or Slavery? The Spanish Crown's Choice of Labor Organization in Sixteenth-Century Spanish America." *The Journal of Economic History* 55(4):842–859.

Zavala, Silvio A. 1973. *La encomienda indiana*. 2 ed. México, DF: Porrúa.