Online Appendix for Trade, Merchants, and the Lost Cities of the Bronze Age (not for publication)

Gojko Barjamovic[*]
Thomas Chaney[†]
Kerem Coşar[‡]
Ali Hortaçsu[§]

Abstract

This appendix contains additional information and robustness checks. It is organized as follows. Section A contains all mathematical proofs and details on our optimization procedure. Section B presents detailed information about our data sources. Section C contains detailed instructions for the construction of optimal travel routes using only topographical data. Section D offers detailed instructions for replicating our analysis of the information contained in merchants' multi-stops itineraries. Section E gives a detailed description of how we construct our $NaturalRoads$ measure. Section F contains additional tables.

[*]Harvard University, barjamovic@fas.harvard.edu.
[†]Sciences Po, thomas.chaney@gmail.com (corresponding author).
[‡]University of Virginia and CEPR, keremcosar@gmail.com.
[§]University of Chicago and NBER, hortacsu@gmail.com.

A. Mathematical Proofs and Optimization

Derivation of shipment probabilities, equations (2), (3), and (4). Equation (2) is the unconditional probability that origin city i is the cheapest source for good ω in destination city j. Given the assumption of Weibull distributed costs in (1), the probability distribution for the cost of delivering good ω from origin i to destination j is also Weibull,

$$\begin{aligned} G_{ij}(c) &= \Pr\left[c_{ij}(\omega) \le c\right] \\ &= \Pr\left[\tau_{ij} c_i(\omega) \le c\right] \\ &= 1 - \exp\left(-T_i \left(w_i \tau_{ij}\right)^{-\theta} c^{\theta}\right). \end{aligned}$$

Equation (2) is then derived exactly as in Eaton and Kortum (2002),

$$\begin{aligned} \Pr\left[c_{ij}(\omega) \le \min_k \{c_{kj}(\omega)\}\right] &= \Pr\left[c_{ij}(\omega) \le \min_{k \ne i} \{c_{kj}(\omega)\}\right] \\ &= \int_0^{\infty} \Pi_{k \ne i} \left(1 - G_{kj}(c)\right) dG_{ij}(c). \end{aligned}$$

We use the c.d.f. $G_{kj}(c) = 1 - \exp\left(-T_k \left(\tau_{kj} w_k\right)^{-\theta} c^{\theta}\right)$ and the corresponding p.d.f. $dG_{ij}(c) = \theta T_i \left(\tau_{ij} w_i\right)^{-\theta} c^{\theta-1} \exp\left(-T_i \left(\tau_{ij} w_i\right)^{-\theta} c^{\theta}\right) dc$ to get,

$$\begin{aligned} \Pr\left[c_{ij}(\omega) \le \min_k \{c_{kj}(\omega)\}\right] &= T_i \left(\tau_{ij} w_i\right)^{-\theta} \int_0^{\infty} \Pi_k \exp\left(-T_k \left(\tau_{kj} w_k\right)^{-\theta} c^{\theta}\right) \theta c^{\theta-1} dc \\ &= T_i \left(\tau_{ij} w_i\right)^{-\theta} \int_0^{\infty} \exp\left(-\left(\sum_k T_k \left(\tau_{kj} w_k\right)^{-\theta}\right) c^{\theta}\right) \theta c^{\theta-1} dc \\ &= T_i \left(\tau_{ij} w_i\right)^{-\theta} \left[\frac{-\exp\left(-\left(\sum_k T_k \left(\tau_{kj} w_k\right)^{-\theta}\right) c^{\theta}\right)}{\sum_k T_k \left(\tau_{kj} w_k\right)^{-\theta}}\right]_0^{\infty} \\ &= \frac{T_i \left(\tau_{ij} w_i\right)^{-\theta}}{\sum_k T_k \left(\tau_{kj} w_k\right)^{-\theta}}. \square \end{aligned}$$

Equation (3) is the probability that origin i is the cheapest source for good ω in destination j, conditional on j not sourcing good ω internally,

$$\Pr\left[c_{ij}(\omega) \le \min_{k \ne j} \{c_{kj}(\omega)\} \,|\, c_{jj}(\omega) > \min_{k \ne j} \{c_{kj}(\omega)\}\right] = \frac{\Pr\left[c_{ij}(\omega) \le \min_k \{c_{kj}(\omega)\}\right]}{\Pr\left[c_{jj}(\omega) > \min_{k \ne j} \{c_{kj}(\omega)\}\right]}.$$

The numerator is given above (derivation of (2). For the denominator, following similar derivations,

$$\begin{aligned} \Pr\left[c_{jj}(\omega) > \min_{k \ne j} \{c_{kj}(\omega)\}\right] &= 1 - \Pr\left[c_{jj}(\omega) \le \min_{k \ne j} \{c_{kj}(\omega)\}\right] \\ &= 1 - \frac{T_j \left(\tau_{jj} w_j\right)^{-\theta}}{\sum_k T_k \left(\tau_{kj} w_k\right)^{-\theta}} \\ &= \frac{\sum_{k \ne j} T_k \left(\tau_{kj} w_k\right)^{-\theta}}{\sum_k T_k \left(\tau_{kj} w_k\right)^{-\theta}}. \end{aligned}$$

Taking the ratio to form a conditional probability, we have the proposed,

$$\Pr\left[c_{ij}(\omega) \leq \min_{k\neq j}\{c_{kj}(\omega)\} \,|c_{jj}(\omega) > \min_{k\neq j}\{c_{kj}(\omega)\}\right] = \frac{T_i(\tau_{ij}w_i)^{-\theta}}{\sum_{k\neq j} T_k(\tau_{kj}w_k)^{-\theta}}. \square$$

Equation (4) is the probability that destination j imports a given good ω from origin i conditional on (a) j not importing from any lost city and (b) j not purchasing good ω internally. Conditions (a) and (b) are satisfied if and only if $\min_{l\in\mathcal{L}\cup\{j\}}\{c_{lj}(\omega)\} > \min_{k\in\mathcal{K}\backslash\{j\}}\{c_{kj}(\omega)\}$. We first characterize the distribution of $\min_{l\in\mathcal{L}\cup\{j\}}\{c_{lj}(\omega)\}$,

$$\begin{aligned}
\Pr\left[\min_{l\in\mathcal{L}\cup\{j\}} c_{lj}(\omega) \leq c\right] &= 1 - \Pr\left[\min_{l\in\mathcal{L}\cup\{j\}} c_{lj}(\omega) > c\right] \\
&= 1 - \Pi_{l\in\mathcal{L}\cup\{j\}} \Pr\left[c_{lj}(\omega) > c\right] \\
&= 1 - \Pi_{l\in\mathcal{L}\cup\{j\}} \left(1 - \Pr\left[c_{lj}(\omega) \leq c\right]\right) \\
&= 1 - \Pi_{l\in\mathcal{L}\cup\{j\}} \exp\left(-T_l(w_l\tau_{lj})^{-\theta} c^{\theta}\right) \\
&= 1 - \exp\left(-\left(\sum_{l\in\mathcal{L}\cup\{j\}} T_l(w_l\tau_{lj})^{-\theta}\right) c^{\theta}\right),
\end{aligned}$$

i.e. a Weibull distribution with shape parameter $\sum_{l\in\mathcal{L}\cup\{j\}} T_l(w_l\tau_{lj})^{-\theta}$. Given this distribution, we can easily form the conditional probability (4), following the same steps as above,

$$\begin{aligned}
&\Pr\left[c_{ij}(\omega) \leq \min_{k\in\mathcal{K}\backslash\{j\}}\{c_{kj}(\omega)\} \,\middle|\, \min_{l\in\mathcal{L}\cup\{j\}} c_{lj}(\omega) > \min_{k\in\mathcal{K}\backslash\{j\}}\{c_{kj}(\omega)\}\right] \\
&= \frac{\Pr\left[c_{ij}(\omega) \leq \min_{k\in\mathcal{K}\cup\mathcal{L}}\{c_{kj}(\omega)\}\right]}{\Pr\left[\min_{l\in\mathcal{L}\cup\{j\}} c_{lj}(\omega) > \min_{k\in\mathcal{K}\backslash\{j\}}\{c_{kj}(\omega)\}\right]} = \frac{\Pr\left[c_{ij}(\omega) \leq \min_{k\in\mathcal{K}\cup\mathcal{L}}\{c_{kj}(\omega)\}\right]}{\Pr\left[\min_{l\in\mathcal{L}\cup\{j\}} c_{lj}(\omega) > \min_{k\in\mathcal{K}\cup\mathcal{L}}\{c_{kj}(\omega)\}\right]} \\
&= \frac{\frac{T_i(\tau_{ij}w_i)^{-\theta}}{\sum_{k\in\mathcal{K}\cup\mathcal{L}} T_k(\tau_{kj}w_k)^{-\theta}}}{1 - \frac{\sum_{l\in\mathcal{L}\cup\{j\}} T_l(\tau_{lj}w_l)^{-\theta}}{\sum_{k\in\mathcal{K}\cup\mathcal{L}} T_k(\tau_{kj}w_k)^{-\theta}}} = \frac{\frac{T_i(\tau_{ij}w_i)^{-\theta}}{\sum_{k\in\mathcal{K}\cup\mathcal{L}} T_k(\tau_{kj}w_k)^{-\theta}}}{\frac{\sum_{k\in\mathcal{K}\backslash\{j\}} T_k(\tau_{kj}w_k)^{-\theta}}{\sum_{k\in\mathcal{K}\cup\mathcal{L}} T_k(\tau_{kj}w_k)^{-\theta}}} = \frac{T_i(\tau_{ij}w_i)^{-\theta}}{\sum_{k\in\mathcal{K}\backslash\{j\}} T_k(\tau_{kj}w_k)^{-\theta}}. \square
\end{aligned}$$

Derivation of moment condition (7) with multiplicative disturbance term, footnote 10. We follow a lightly edited version of Eaton, Kortum, and Sotelo (2012), and add to the trade cost function a multiplicative disturbance term drawn from a Gamma distribution,

$$\tau_{ij}^{-\theta} = \mu Distance_{ij}^{-\zeta}\nu_{ij}, \text{ with } \nu_{ij} \sim Gamma\left(\frac{1}{\eta^2}\frac{\alpha_i Distance_{ij}^{-\zeta}}{\sum_{k\neq j}\alpha_k Distance_{kj}^{-\zeta}}, \frac{\eta^2}{\alpha_i Distance_{ij}^{-\zeta}}\right).$$

Treating the ν's as realizations from a random variable, we rely on the scaling property of the Gamma distribution to obtain

$$\alpha_i Distance_{ij}^{-\zeta}\nu_{ij} \sim Gamma\left(\frac{1}{\eta^2}\frac{\alpha_i Distance_{ij}^{-\zeta}}{\sum_{k\neq j}\alpha_k Distance_{kj}^{-\zeta}}, \eta^2\right).$$

The joint distribution of n Gamma distributed variables normalized by their sum is Dirichlet,

$$\left(\ldots, \frac{\alpha_i Distance_{ij}^{-\zeta}\nu_{ij}}{\sum_{k\neq j}\alpha_k Distance_{kj}^{-\zeta}\nu_{kj}}, \ldots\right)_{i\neq j} \sim Dirichlet\left(\ldots, \frac{1}{\eta^2}\frac{\alpha_i Distance_{ij}^{-\zeta}}{\sum_{k\neq j}\alpha_k Distance_{kj}^{-\zeta}}, \ldots\right).$$

Note that by definition of the Dirichlet joint distribution, all shares are in $(0,1)$ and they add up to one. Using the definition for the mean of a Dirichlet distribution, we recover our proposed moment condition (7),

$$\mathbb{E}\left[\frac{\alpha_i Distance_{ij}^{-\zeta}\nu_{ij}}{\sum_{k\neq j}\alpha_k Distance_{kj}^{-\zeta}\nu_{kj}}\right] = \frac{\frac{1}{\eta^2}\frac{\alpha_i Distance_{ij}^{-\zeta}}{\sum_{k\neq j}\alpha_k Distance_{kj}^{-\zeta}}}{\sum_{l\neq j}\frac{1}{\eta^2}\frac{\alpha_l Distance_{lj}^{-\zeta}}{\sum_{k\neq j}\alpha_k Distance_{kj}^{-\zeta}}} = \frac{\alpha_i Distance_{ij}^{-\zeta}}{\sum_{k\neq j}\alpha_k Distance_{kj}^{-\zeta}}. \square$$

Derivation of our measure of size, equation (10). To express $Size_i \propto Pop_i T_i^{1/\theta}$ as a function of observables and model parameters only, we first use the definition of our exporter fixed effect α_i estimated from (8),

$$\alpha_i \propto T_i w_i^{-\theta} \Rightarrow Pop_i T_i^{1/\theta} \propto \alpha_i^{1/\theta} w_i Pop_i.$$

From market clearing,

$$w_i Pop_i = X_i \Rightarrow Pop_i T_i^{1/\theta} \propto \alpha_i^{1/\theta} X_i.$$

The volume of trade from i to j is simply equal to total expenditure in j multiplied by the probability of sourcing a good from origin i.

$$X_{ij} = \frac{T_i w_i^{-\theta} Distance_{ij}^{-\theta} X_j}{\sum_k T_k w_k^{-\theta} Distance_{kj}^{-\theta}}.$$

We then manipulate this expression as in Anderson and van Wincoop (2003) to obtain

$$X_{ij} = \frac{T_i w_i^{-\theta}\tau_{ij}^{-\theta} X_j}{\sum_k T_k w_k^{-\theta}\tau_{kj}^{-\theta}} = \frac{X_i X_j}{X_{total}}\left(\frac{\tau_{ij}}{\Pi_i P_j}\right)^{-\theta},$$

with $\Pi_i^{-\theta} = \sum_k \left(\frac{\tau_{ik}}{P_k}\right)^{-\theta}\frac{X_k}{X_{total}}$ a measure of outward resistance, $P_j^{-\theta} = \sum_k \left(\frac{\tau_{kj}}{\Pi_k}\right)^{-\theta}\frac{X_k}{X_{total}}$ a measure of inward resistance, and $X_{total} = \sum_k X_k$. We will rely on the result that if trade frictions are symmetric, $\tau_{ij} = \tau_{ji}, \forall i \neq j$, then $\Pi_i = P_j$ and expected trade is symmetric, $X_{ij} = X_{ji}$. Using the equivalence between trade shares in value and in count in the Eaton and Kortum (2002) model,

$$\frac{X_{ij}}{X_j} = \frac{T_i w_i^{-\theta}\tau_{ij}^{-\theta}}{\sum_k T_k w_k^{-\theta}\tau_{kj}^{-\theta}} = \mathbb{E}\left[\frac{N_{ij}}{\sum_k N_{kj}}\right] = \frac{\alpha_i \tau_{ij}^{-\theta}}{\sum_k \alpha_k \tau_{kj}^{-\theta}}.$$

Combining this with the above expression for bilateral trade, we get

$$\frac{X_{ij}}{X_j} = \frac{X_i}{X_{total}}\left(\frac{\tau_{ij}}{\Pi_i P_j}\right)^{-\theta} = \frac{\alpha_i \tau_{ij}^{-\theta}}{\sum_k \alpha_k \tau_{kj}^{-\theta}}, \forall i \neq j \Rightarrow X_i \propto \alpha_i \Pi_i^{-\theta}.$$

From the above, the definition of $\Pi_i^{-\theta}$, and symmetry, $P_k = \Pi_k$, we derive

$$\Pi_i^{-\theta} \propto \sum_k \tau_{ik}^{-\theta} X_k / P_k^{-\theta} = \sum_k \tau_{ik}^{-\theta} X_k / \Pi_k^{-\theta} \propto \sum_k \tau_{ik}^{-\theta}\alpha_k.$$

Combining $\tau_{ik}^{-\theta} \propto Distance_{ik}^{-\zeta}$ and the above, we get the proposed formula,

$$Size_i \ \propto \ Pop_i T_i^{1/\theta} \ \propto \ \alpha_i^{1+1/\theta}\sum_k Distance_{ki}^{-\zeta}\alpha_k. \square$$

Analytical formulas for standard errors. We follow Cameron and Trivedi (2005) throughout. The notation is also borrowed from that book, specifically pages 200-202. Let $\boldsymbol{\theta} = (\zeta, \cdots)'$ be the parameters of the first step (PPML), collecting the distance elasticity of trade ζ and the exporter and importer fixed effects. $\boldsymbol{\beta}$ is the vector of parameters of the second step (NLLS). Let K_1 be the number of city pairs for step 1, and K_2 the number of city pairs for step 2. K_1 includes all known cities in our sample that import from other known cities that import from other known cities ... etc. K_2 includes all cities (known or lost) in our sample, so that $K_1 < K_2$.

The variance-covariance matrix of $(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\beta}})$ is

$$\hat{\boldsymbol{\Sigma}} = \widehat{var}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\beta}}) = \frac{1}{K_2} \begin{pmatrix} G_{11} & 0 \\ G_{21} & G_{22} \end{pmatrix}^{-1} \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} G_{11} & 0 \\ G_{21} & G_{22} \end{pmatrix}^{-1\top}.$$

We now explain what the components of these block matrices are.

Consider first the Poisson Pseudo MLE step. Let y_{ij} be the outcome variable, trade share in counts, and x_{ij} be the vector of covariates. The first component of $\boldsymbol{\theta}$ is the distance elasticity of trade. The other components are coefficients associated with the exporter and importer dummies. The log-likelihood is then:

$$\log \mathcal{L}(\boldsymbol{\theta}) = \sum_{i,j} \left(-\exp\{x'_{ij}\boldsymbol{\theta}\} + y_{ij} x'_{ij}\boldsymbol{\theta} - \log y_{ij}! \right).$$

In m-Estimators terminology (Cameron and Trivedi, p.118), we have:

$$Q_{K_1}(\boldsymbol{\theta}) = \frac{1}{K_1} \sum_{i,j} \left(-\exp\{x'_{ij}\boldsymbol{\theta}\} + y_{ij} x'_{ij}\boldsymbol{\theta} - \log y_{ij}! \right) = \frac{1}{K_1} \sum_{i,j} q_{i,j}(\boldsymbol{\theta}).$$

In the notation of two-step m-Estimation (p.200), we have that $h_1(w_{ij}, \boldsymbol{\theta}) = \frac{d}{d\boldsymbol{\theta}} q_{ij}(\boldsymbol{\theta})$. Thus,

$$G_{11} = \frac{1}{K_1} \sum_{i,j} \frac{dh_1(w_{ij}, \hat{\boldsymbol{\theta}})}{d\boldsymbol{\theta}'} = \frac{1}{K_1} \sum_{i,j} \frac{d^2 q_{ij}(\hat{\boldsymbol{\theta}})}{d\boldsymbol{\theta} d\boldsymbol{\theta}'} = \frac{1}{K_1} \sum_{i,j} \exp\{x'_{ij}\boldsymbol{\theta}\} x_{ij} x'_{ij}$$

and

$$S_{11} = \frac{1}{K_1} \sum_{ij} h_1(w_{ij}, \hat{\boldsymbol{\theta}}) h_1(w_{ij}, \hat{\boldsymbol{\theta}})' = \frac{1}{K_1} \sum_{ij} \frac{dq_{ij}(\hat{\boldsymbol{\theta}})}{d\theta_1} \frac{dq_{ij}(\hat{\boldsymbol{\theta}})}{d\theta'_1} = \frac{1}{K_1} \sum_{ij} (y_{ij} - \exp\{x'_{ij}\boldsymbol{\theta}\})^2 x_{ij} x'_{ij}.$$

This coincides with the BHHH estimate (p.138).

Consider now the standard non-linear least squares step. Let e_{ij} be the difference between the model and data trade shares. Given the first stage estimate $\hat{\boldsymbol{\theta}}$, the first order condition of the least squares problem is:

$$\sum_{i,j} e_{ij}(\hat{\boldsymbol{\beta}}; \hat{\boldsymbol{\theta}}) \frac{\partial}{\partial \boldsymbol{\beta}} e_{ij}(\hat{\boldsymbol{\beta}}; \hat{\boldsymbol{\theta}}) = 0.$$

In the notation of the two-step m-Estimation (p.200), we have that $h_2(w_{ij}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\beta}}) = e_{ij}(\hat{\boldsymbol{\beta}}; \hat{\boldsymbol{\theta}}) \frac{\partial}{\partial \boldsymbol{\beta}} e_{ij}(\hat{\boldsymbol{\beta}}; \hat{\boldsymbol{\theta}})$. The nonlinear least squares simplifications from p.153 apply, so that

$$G_{22} = \frac{1}{K_2} \sum_{i,j} \frac{\partial}{\partial \boldsymbol{\beta}} e_{ij}(\hat{\boldsymbol{\beta}}; \hat{\boldsymbol{\theta}}) \frac{\partial}{\partial \boldsymbol{\beta}'} e_{ij}(\hat{\boldsymbol{\beta}}; \hat{\boldsymbol{\theta}}).$$

Following p.201, we have

$$S_{22} = \frac{1}{K_2}\sum_{i,j} h_{2i}h'_{2i} = \frac{1}{K_2}\sum_{i,j}\left[e_{ij}(\hat{\boldsymbol{\beta}};\hat{\boldsymbol{\theta}})\frac{\partial}{\partial\boldsymbol{\beta}}e_{ij}(\hat{\boldsymbol{\beta}};\hat{\boldsymbol{\theta}})\right]\left[e_{ij}(\hat{\boldsymbol{\beta}};\hat{\boldsymbol{\theta}})\frac{\partial}{\partial\boldsymbol{\beta}'}e_{ij}(\hat{\boldsymbol{\beta}};\hat{\boldsymbol{\theta}})\right]$$

Finally, we compute the interactions terms, G_{21}, S_{12} and S_{21}. Given that we know h_1 and h_2, we use the following variation on p.202:

$$\begin{aligned} G_{21} &= \frac{1}{K_2}\sum_{i,j}\frac{\partial h_2}{\partial\boldsymbol{\theta}'} = \frac{1}{K_2}\sum_{i,j}\frac{\partial}{\partial\boldsymbol{\theta}'}\left[e_{ij}(\hat{\boldsymbol{\beta}};\hat{\boldsymbol{\theta}})\frac{\partial}{\partial\boldsymbol{\beta}}e_{ij}(\hat{\boldsymbol{\beta}};\hat{\boldsymbol{\theta}})\right] \\ S_{12} &= \frac{1}{K_1}\sum_{i,j} h_{1i}h'_{2i} = \frac{1}{K_1}\sum_{i,j}\left(\frac{d}{d\boldsymbol{\theta}}q_{ij}(\hat{\boldsymbol{\theta}})\right)e_{ij}(\hat{\boldsymbol{\beta}};\hat{\boldsymbol{\theta}})\frac{\partial}{\partial\boldsymbol{\beta}'}e_{ij}(\hat{\boldsymbol{\beta}};\hat{\boldsymbol{\theta}}) \\ S_{21} &= S_{12}^{\top}. \end{aligned}$$

To derive the standard errors for city sizes, $Size_i$, we apply the Delta method to (10), using the above covariance matrix $\hat{\boldsymbol{\Sigma}}$.

Analytical formulas for iso-density contours and** precision (l) **in (12). Iso-density contours are points with latitude-longitude (φ, λ) such that $f_l(\varphi, \lambda) = c$ where f_l is the p.d.f. of the bi-variate normal distribution with estimated mean $(\hat{\varphi}_l, \hat{\lambda}_l)$, variance $(\hat{\sigma}^2_{\varphi_l}, \hat{\sigma}^2_{\lambda_l})$, and correlation $\hat{\rho}_{\varphi_l,\lambda_l}$,

$$c = \frac{1}{2\pi\hat{\sigma}_{\varphi_l}\hat{\sigma}_{\lambda_l}\sqrt{1-\hat{\rho}^2_{\varphi_l,\lambda_l}}}\exp\left(-\frac{1}{2(1-\hat{\rho}^2_{\varphi_l,\lambda_l})}\left[\frac{(\varphi-\hat{\varphi}_l)^2}{\hat{\sigma}^2_{\varphi_l}} + \frac{(\lambda-\hat{\lambda}_l)^2}{\hat{\sigma}^2_{\lambda_l}} - \frac{2\hat{\rho}_{\varphi_l,\lambda_l}(\varphi-\hat{\varphi}_l)(\lambda-\hat{\lambda}_l)}{\hat{\sigma}_{\varphi_l}\hat{\sigma}_{\lambda_l}}\right]\right).$$

This is the formula for an ellipse. We use four values for c corresponding to 50%, 75%, 90%, and 99% confidence regions. For instance, for the 75th percentile, we use c_{75} defined as,

$$\Pr\left[(\varphi,\lambda)\mid \text{s.t. } f_l(\varphi,\lambda) \geq c_{75}\right] = 0.75.$$

To derive an analytical formula for *precision*(l), we start from the definition,

$$precision\,(l) = \sqrt{\mathbb{E}_{(\varphi,\lambda)\sim\mathcal{N}(\hat{\boldsymbol{\beta}}_l,\hat{\boldsymbol{\Sigma}}_l)}\left[\left(Distance\left(\hat{\varphi}_l,\hat{\lambda}_l;\varphi,\lambda\right)\right)^2\right]}.$$

Using the Euclidean formula for distance and linearity of the expectation operator, we get,

$$\begin{aligned} \mathbb{E}_{(\varphi,\lambda)\sim\mathcal{N}(\hat{\boldsymbol{\beta}}_l,\hat{\boldsymbol{\Sigma}}_l)}\left[\left(Distance\left(\hat{\varphi}_l,\hat{\lambda}_l;\varphi,\lambda\right)\right)^2\right] &= \left(\frac{10000}{90}\right)^2\left(\mathbb{E}\left[(\varphi-\hat{\varphi}_l)^2\right] + \cos^2\left(\frac{37.9}{180}\pi\right)\mathbb{E}\left[(\lambda-\hat{\lambda}_l)^2\right]\right) \\ &= \left(\frac{10000}{90}\right)^2\left[\hat{\sigma}^2_{\varphi_l} + \cos^2\left(\frac{37.9}{180}\pi\right)\hat{\sigma}^2_{\lambda_l}\right] \end{aligned}$$

Thus, our formula for the geographic precision (in kms) for lost city l is,

$$precision\,(l) = \frac{10000}{90}\sqrt{\hat{\sigma}^2_{\varphi_l} + \cos^2\left(\frac{37.9}{180}\pi\right)\hat{\sigma}^2_{\lambda_l}},$$

where $(\hat{\sigma}^2_{\varphi_l}, \hat{\sigma}^2_{\lambda_l})$ are the variances for the estimated latitudes and longitudes $(\hat{\varphi}_l, \hat{\lambda}_l)$ of city l.

Summary of our Optimization Procedure. Our estimates of $(\zeta; \cdots (\varphi_l, \lambda_l) \cdots ; \cdots \alpha_i \cdots)$ are computed in two steps. First, the distance elasticity of trade – ζ – is estimated by Poisson Pseudo Maximum Likelihood in the subsample with known location data. Specifically, the independent variable in the Poisson regression model is the observed trade shares, and the dependent variables are the log of distance between cities, destination city dummies and origin city dummies. The estimate for ζ is the resulting coefficient for the log of distance between cities. Our estimation uses the `ppml` STATA command written by Silva and Tenreyro (2006).

Second, the geo-coordinates of lost cities and the α_i's are estimated by minimizing the sum of squared differences between observed and predicted trade shares given our estimate of ζ. This function of our parameters was coded in Python-Numpy, and the optimization performed using IPOPT. In particular, we ran 260 jobs in parallel. Each job executed the minimization process 20 times, starting from random initial values. The geo-coordinates' initial values were uniformly drawn between 36 and 42 degrees of latitude and 27 and 45 degrees of longitude, whereas the α_i's initial values were uniformly drawn between 0 and 200. The IPOPT specifications include: $100,000$ maximum iterations, `MA57` as the linear solver, a tolerance level of $1.0e^{-08}$, and an acceptable tolerance level of $1.0e^{-07}$. Issues of local minima are dealt away upon inspection of all minimization results. For our main specification, 242 out of $5,200$ executions yield the same lowest sum of squared differences and have the same parameter estimates, up to 6 decimal points.

B. Data Sources and Optimization Procedure

B.1. Data Construction and Selection of Cities

The database of ancient texts that we have access to contains references to 79 unique cities in Anatolia (Barjamovic, 2011). Decades of scholarship has successfully disambiguated place names from terms previously mistaken as cities while they actually refer to types of textiles or people.[1] We drop 40 cities with only a single mention in the entire corpus: their parameters are not identified by our structural gravity estimation. Other than the remaining 39 Anatolian cities that appear more than once in the most up to date corpus of Assyrian texts, there are two cities *outside of Anatolia* with known locations: *Aššur* and *Qattara*. As explained in the text (see footnote 3), we exclude *Aššur* from our analysis because the word for *Aššur* is ambiguous, appearing in persons' names and as the name of the main Assyrian deity. This makes it impossible to run an automated search for the city in the digitized text. We also drop the city of *Qattara* because it was known to be a small independent polity in proximity to *Aššur*, where Assyrian merchants were allowed to pass by, but not to trade. *Qattara* being also the first stop on the way from *Aššur* to Anatolia (last stop on return trips), any mention of *Qattara* can only mean a shipment to or from *Aššur*, not a shipment to or from *Qattara* itself. Similar to *Aššur*, this city is also outside of the modern day Turkish geography we focus on.

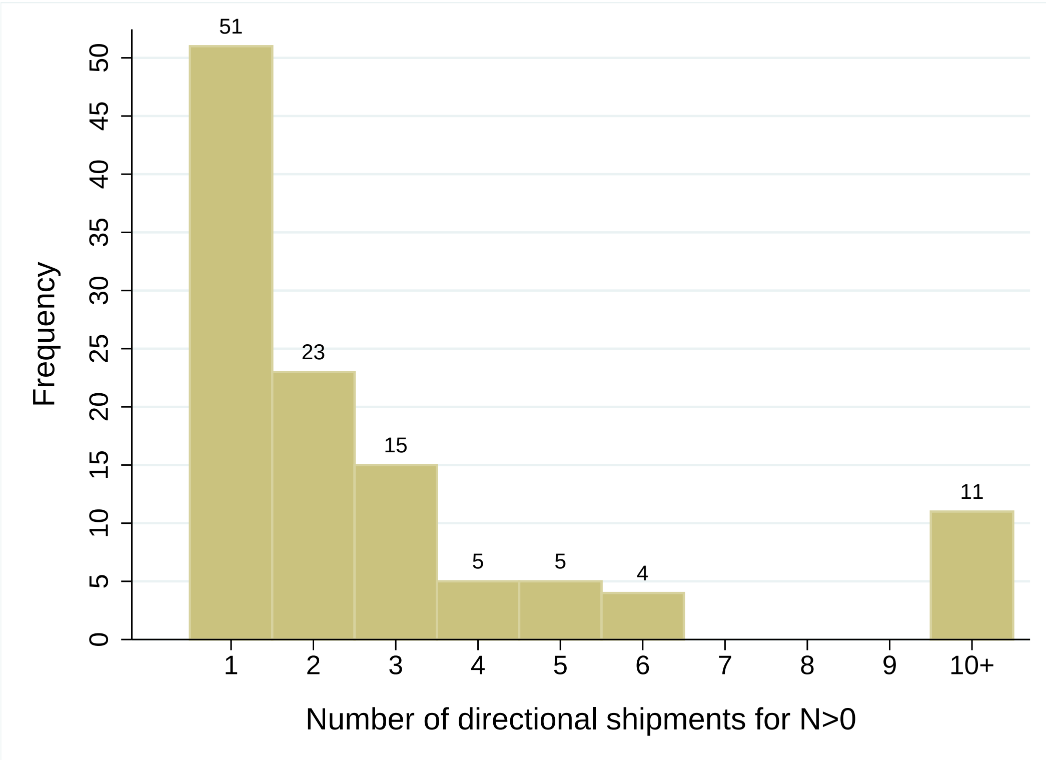
The second criterion is that, out of the set of 39 Anatolian cities, we drop 9 lost cities about which professional Assyriologists have only speculative conjectures in terms of locations. This is due to both the relatively low number of mentions to these cities in the texts as well as the lack of precise descriptions of their geography—see Barjamovic (2011), Table 39 in page 411.

The final criterion is mechanical: any city that is disconnected from the rest of the network of trading cities cannot be identified in a gravity model. If there is no trade at all between city i and any other city, our estimator will assign a size zero to that city, and/or an infinite distance if this city is lost. Note that network connectedness is an iterative notion. We start by dropping

[1] For more details, see Bilgiç (1951) and Michel and Veenhof (2010).

all disconnected cities. Having dropped those cities, some cities which were only trading with the dropped cities are dropped in turn, which further eliminates the cities that were only trading with those, and so forth. This iterative process leaves us with 25 cities.

The next figure shows the distribution of shipment counts, for all (directed) city pairs with a positive number of shipments, $N_{ij}^{data} > 0$.



Appendix Figure 1: **Frequency of shipment counts.**

B.2. An Example of a Coincidental Joint Attestation of Two Cities

After reading the 2,806 tablets which mention at least two cities, we discard any case where the two cities are mentioned in the same tablet for reasons unrelated to any trade relationship between those cities. Below is an example of a purely coincidental joint attestation of two city names, *Hahhum* and *Wahšušana*, underlined in the text for clarity,

> *From Enlil-bani to Aššur-idi with Cc to Ṭab-ṣilli-Aššur and Aššur-[xx] . We had half a pound of silver transferred in Hahhum, and I gave it to you in the year of Aššur-malik. You said: "After I make a transport to the City, I will either give you the proceeds or I will take goods on commission for you." I paid you in refined silver and we saw each other about five times when you cheated me. While I was staying in Wahšušana, my representatives seized you, but you gave them nothing. My representatives said ...*
>
> [Tablet BIN 6, 38 (NBC 3808) lines 1-18]

In this text, the merchant Enlil-bani accuses another merchant, Ṭab-ṣilli-Aššur, of having cheated him. Enlil-bani mentions the city of *Hahhum* in the context of a financial transaction between both merchants (second line). The city of *Wahšušana* is mentioned in passing in the same letter (penultimate line) to describe Enlil-bani's whereabouts while his representatives were trying to recover his due. It has nothing to do with the initial financial transaction. There is no direct economic connection between both cities that can be drawn from this letter.

B.3. A Partial Example of how Historians Locate Lost Cities

Historians Forlanini (2008) and Barjamovic (2011) use a series of references to ancients cities during the Middle Bronze Age period, complemented with references from later periods, to make informed proposals for the location of lost cities. A detailed list of references to such proposals was collected by Nashef (1992) with updates in Ullmann and Weeden (2017).

An example of a few snippets of information that guide the reasoning of historians would be the following references related to the city of *Hahhum*. The following text locates *Hahhum* on a river:

> *I met Elali in Hahhum while I was staying there at the bank of the river in Habnuk.* [Tablet I 469]

Another text, in this case a later Hittite royal annal, makes it explicit that the river in question was the Euphrates:

> *I the Great King Tabarna took away from Hahhu and presented it to the Sun-God. The Great King Tabarna removed the hands of its slave girls from the grindstone and its slaves? hands he removed ... he released their belts and he put them in the temple of the Sun-Goddess of Arinna ... The great river Euphrates, no one had crossed it. [The Great King] Tabarna crosses it on foot, and his troops crossed it [on] foot after him*
>
> [Text KBo 10.1]

A large number of additional references eventually allows historians to formulate with confidence the hypothesis that *Hahhum* lies to the South and East of *Kaneš*, in a position that allowed it to control an important river crossing. Its neighbor on the opposite river bank was *Badna*, and *Timelkiya* was the next state on the route north-west to *Kaneš*. Gradually, the positions of the cities can be established in relation to one another to form a network that can then be placed on a map. In rare cases, place names may survive to help guide this process, if their location in later times is known.

B.4. Data on Modern-day Trade, Local Resources, and Topography

Modern-day population. Data on the 2012 urban population of Turkish districts is obtained from the website of the Turkish Statistical Agency (https://biruni.tuik.gov.tr/EdUygulamaDis/zul/loginEN.zul?lang=en).

Night time luminosity. Data on nighttime light emissions intensity is used as a proxy for local GDP at the very granular level, as in Hodler and Raschky (2014). The data for 2003 is from the National Oceanic and Atmospheric Administration (NOAA), available at http://ngdc.noaa.gov. Weather satellites from the U.S. Air Force measure light intensity between 8:30PM and 10:00PM, removing observations affected by cloud conditions, and correcting for likely ephemeral lights or background noise.

Crop yields. We use the low-input level rain-fed cereal suitability index of IIASA/FAO (2012) available at http://www.fao.org/nr/gaez/en/.

Ruggedness. We use the Terrain Ruggedness Index of Riley, DeGloria, and Elliot (1999), defined as the square root of the sum of squared elevation changes in 8 directions.

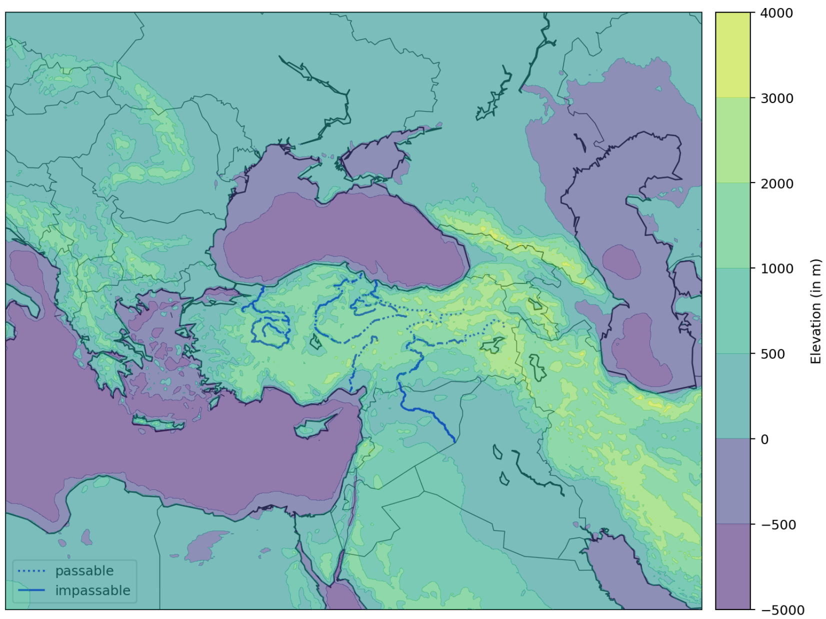
Elevation. Elevation data to calculate the natural road scores is obtained from IIASA/FAO (2012) and is available at `http://www.fao.org/nr/gaez/en/`.

Rivers and lakes. The shapefile of rivers and lakes in Turkey, used in calculating the natural road scores, has been downloaded from http://www.naturalearthdata.com/.

C. Optimal Travel Routes

To define optimal travel routes, we use topographical data, and Dijkstra's shortest path algorithm.

Formally, we collect data on elevation on a fine grid (each pixel's side is 5 arc minutes, or about 10 km). The elevation data is downloaded from FAO-GAEZ, which itself is based on NASA's shuttle radar topography mission (IIASA/FAO, 2012). We collect elevation data on a large area around central Anatolia, in order to avoid a core-periphery bias –the tendency to have more road-crossings in locations in the center of the map. The total area is contained between 16 and 55 degrees of longitude East, and 26 and 51 degrees of latitude North. It corresponds to a wide area between Hungary in the Northwest, Kazakhstan in the Northeast, Kuwait in the Southeast, and Libya in the Southwest, which extends well beyond central Anatolia. We remove from this map the Arabian desert, assuming implicitly that travellers were not crossing it. We do include any maritime area where the sea is less than 500 m deep, allowing maritime travel along the coasts, but preventing high sea travel. The elevation map for the entire region we consider is depicted below.



Appendix Figure 2: **Elevation.**

Given this fine elevation grid, we compute travel times between any pixel and its eight neighbors (North-South, East-West, and diagonally). First, we compute the horizontal distance between any two contiguous pixels (in meters), and the signed elevation difference between them (in meters).

We then apply the formula from Langmuir (1984) to translate distances and slope into travel times (in seconds). The parametrization of Langmuir's formula we apply for overland travel is as follows. We assume it takes 0.72 seconds to travel 1 meter horizontally; it takes an additional 6 seconds for each vertical meter uphill; going downhill 1 vertical meter on a gentle slope (less than or equal to 21.25%) saves 2 seconds per vertical meter; going downhill on a steep slope (more than 21.25%) adds an additional 2 seconds per vertical meter. For maritime travel, we assume traveling by boat is 10% faster than traveling over a featureless plain overland. However, in order to avoid many very short trips by sea, we assume it takes 1 hour to embark on a boat (from land to sea), and 1 hour to disembark (from sea to land). This assumption of a fixed cost of loading/unloading boats creates a natural tendency for sea ports to emerge where natural terrestrial routes (e.g. a valley) connect to the sea. Finally, we manually code major lakes, and three main rivers, the Euphrates, the Red river, and the Green river in Turkey. For the three rivers, we collect information on impassable segments of the river (e.g. deep and steep canyon), as well as easy crossings/fords known to have been used in the Bronze Age, using data from Palmisano (2013). We impose a prohibitive penalty for crossing major lakes and those three rivers over segments where they are deemed impassable, and allow crossing as if it were on dry land for the river crossings.

Having defined travel times between any pixel and its eight neighbors, we apply Dijkstra's algorithm to compute the optimal travel paths between any two pixels (Dijkstra, 1959). We use those travel paths and travel times when encoding the information contained in merchants itineraries in section IV.B (see the details in appendix D), and when defining natural roadways to compute the variable $NaturalRoads$ in section V.B (see the details in appendix E).

D. Constraints from Merchants Itineraries

To impose constraints on the location of lost cities, using information contained in multi-stop merchant itineraries, we proceed as follows.

First, we collect systematic information on multiple itineraries of either merchants or caravans in our corpus of texts. We keep only itineraries with at least one stop in a lost city.

Second, we compute the following two statistics for all segments of those itineraries from one known city to another known city: the average travel length $||\text{average segment}||$, and the standard deviation of the segment lengths $||\text{s.d. segment}||$. Our measure of length is the optimal travel time between the two ends of each segment, defined in appendix C.

Third, we jointly impose on all lost cities the "short detour" and "pit stop" constraints defined in section IV.B, using all mentions of itineraries. In other words, for all lost cities jointly, we search for all grid points such that both constraints are satisfied. To solve for this multi-dimensional search, we proceed sequentially. We start by imposing all "pit stop" constraints where one city is known, and one city is unknown. This gives us an admissible region for each lost city mentioned at least once alongside a known city. We then impose all the "short detour" constraints involving two known cities and one lost city, searching only within the admissible regions of the previous step. This further restricts the size of the admissible regions for each lost city mentioned at least once alongside known cities. We finally solve a minimization problem using all the remaining "pit stop" and "short detour" constraints, imposing a penalty for a violation of the constraints.

E. Constructing Natural Road Scores

To compute the $NaturalRoads$ variable, we proceed as follows.

We use the large region depicted in Appendix Figure 2, which extends widely beyond central Anatolia. For any two pixels on that map, we know the route of the shortest path from one to the other, using the procedure described in appendix C. Our purpose is to compute, for each pixel on the map, the number of optimal paths that intersect on that pixel. This corresponds to the notion of between-ness centrality in the network of optimal routes.

In order to distinguish between short distance routes (arguably travelled frequently) versus long distance routes (arguably travelled infrequently), we implicitly assume a gravity model of migrations. For any optimal route from pixel i to pixel j with duration d_{ij}, we assume this route is travelled with a probability proportional to $d_{ij}^{-\hat{\zeta}}$, where $\hat{\zeta} = 1.9$ is our estimate for the distance elasticity of trade in the Middle Bronze Age, and where we use the algorithm described in appendix C to compute shortest paths and their durations. This probability weighting corresponds to an implicit gravity model of migrations, where humans are uniformly distributed over space, and travel between locations according to a gravity model with distance elasticity $\hat{\zeta}$.

For any pair of pairs of points on Appendix Figure 2, (A,B) and (C,D), with shortest paths of durations d_{AB} and d_{CD}, we draw the pair (A,B) with probability proportional to $d_{AB}^{-\hat{\zeta}}$ and the pair (CD) with probability proportional to $d_{CD}^{-\hat{\zeta}}$. If the two paths either intersect or overlap on a given pixel, we record their intersection for that pixel point. We repeat this procedure 1 million times. Each pixel receives a "road-knot score" equal to the number of intersections or overlaps recorded on that pixel.

For each ancient city i, either known or lost (we use our structural gravity estimates for the location of lost cities as our main specification, but also experiment with alternative locations as robustness), our variable of interest, $NaturalRoads_i$, simply adds up this "road-knot score" from all pixels which are within 20 km of city i.

References

Anderson, J. E. and E. van Wincoop (2003): "Gravity with Gravitas: A Solution to the Border Puzzle," *American Economic Review*, 93, 170–92.

Barjamovic, G. (2011): *A Historical Geography of Anatolia in the Old Assyrian Colony Period*, Copenhagen: Museum Tusculanum Press.

Bilgiç, E. (1951): "Die Ortsnamen der kappadokischen Urkunden im Rahmen der alten Sprachen Anatoliens," *Archiv für Orientforschung*, 15, 3–37.

Cameron, A. C. and P. K. Trivedi (2005): *Microeconometrics: methods and applications*, Cambridge university press.

Dijkstra, E. W. (1959): "A note on two problems in connexion with graphs," *Numerische mathematik*, 1, 269–271.

Eaton, J. and S. Kortum (2002): "Technology, Geography and Trade," *Econometrica*, 70, 1741–79.

Eaton, J., S. Kortum, and S. Sotelo (2012): "International Trade: Linking Micro and Macro," *NBER Working Paper No.17864*.

Forlanini, M. (2008): "The Central Provinces of Hatti. An Updating," in *New Perspectives on the Historical Geography and Topography of Anatolia in the II and I Millennium BC*, ed. by K. Strobel, (EOTHEN 16) Firenze: LoGisma Editore, 1, 145–188.

Hodler, R. and P. A. Raschky (2014): "Regional Favoritism," *The Quarterly Journal of Economics*, 129, 995–1033.

IIASA/FAO (2012): *Global Agro-Ecological Zones (GAEZ v3.0)*, IIASA, Laxenburg, Austria and FAO, Rome.

Langmuir, E. (1984): *Mountaincraft and leadership: a handbook for mountaineers and hillwalking leaders in the British Isles*, Edinburgh: Scottish Sports Council.

Michel, C. and K. Veenhof (2010): "The Textiles Traded by the Assyrians in Anatolia (19th-18th centuries BC)," in *Textile Terminologies in the Ancient Near East and Mediterranean from the Third to the First Millennia BC*, ed. by C. Michel and M.-L. Nosch, Ancient Textiles Series Vol. 8. Oxford and Oakville: Oxbow Books, 210–271.

Nashef (1992): *Die Orts- und Gewassernamen der altassyrischen Zeit*, Wiesbaden: Ludwig Reichert.

Palmisano, A. (2013): "Computational and Spatial Approaches to the Commercial Landscapes and Political Geography of the Old Assyrian Colony Period." in *Time and History in the Ancient Near East. Proceedings of the 56th Rencontre Assyriologique Internationale, Barcelona, July 26-30, 2010.*, ed. by L. Feliu, J. Llop, A. M. Albà, and W. Lake., Eisenbrauns, 767–783.

Riley, S. J., S. D. DeGloria, and R. Elliot (1999): "Index that quantifies topographic heterogeneity," *Intermountain Journal of sciences*, 5, 23–27.

Silva, J. S. and S. Tenreyro (2006): "The Log of Gravity," *The Review of Economics and statistics*, 88, 641–658.

Tobler, W. and S. Wineburg (1971): "A Cappadocian speculation," *Nature*, 231, 39–41.

Ullmann, L. and M. Weeden (2017): *Hittite Landscape and Geography*, Leiden: Brill.

F. Additional Tables

Appendix Table 1: **Proof of Concept, Recovering Fictitiously Lost Cities**

	True Coordinates		Estimated Coordinates		
	Latitude	Longitude	Latitude	Longitude	Distance in km
Hattus	40.021	34.61	40.044	34.635	3
			(0.53)	(0.58)	
Kanes	38.85	35.633	38.955	35.277	33
			(0.407)	(0.375)	
Karahna	40	36.1	40.021	34.618	130
			(2964.306)	(170.666)	
Tapaggas	40.148	35.762	40	35.818	17
			(317.12)	(97.762)	
Hanaknak	40	35.817	40.164	35.764	19
			(0.015)	(0.24)	
Hurama	38.261	37.114	39.643	37.327	155
			(0.715)	(0.457)	
Malitta	39.363	33.787	38.996	34.644	86
			(0.659)	(0.442)	
Mamma	37.583	36.933	38.02	36.5	62
			(1.603)	(0.513)	
Salatuwar	39.655	31.994	39.64	33.192	105
			(0.895)	(0.413)	
Samuha	39.619	36.528	39.399	36.11	44
			(0.196)	(0.383)	
Timelkiya	38.027	38.234	38.403	37.488	78
			(0.451)	(0.411)	
Ulama	38.411	33.834	39.852	33.196	170
			(0.402)	(0.966)	
Unipsum	38.021	36.503	37.583	36.933	61
			(1462.246)	(1105.701)	
Wahsusana	39.584	33.418	38.457	32.077	172
			(17.831)	(27.255)	
Zimishuna	40.461	35.65	40.47	35.65	1
			(307617.802)	(461223.631)	
Mean					76
Median					62

Notes: This table presents the results from our 'proof-of-concept' exercise in section IV.C. For all known cities, we list the true geo-coordinates of each city, their estimated geo-coordinates from estimating a model similar to (8), and the distance, in kms, between the true and estimated locations. The results for all cities are shown on single map on figure VI. All latitudes are North, and all longitudes are East. Robust (White) standard errors in parentheses.

Appendix Table 2: **Proof of Concept (robustness), Recovering One Fictitiously Lost City and Ten Lost Cities Jointly**

	True Coordinates		Estimated Coordinates		
	Latitude	Longitude	Latitude	Longitude	Distance in km
Hattus	40.021	34.61	39.997	36.131	133
			(0.322)	(0.073)	
Kanes	38.85	35.633	39.313	33.918	159
			(0.252)	(0.156)	
Karahna	40	36.1	40.046	34.547	136
			(0.754)	(0.475)	
Tapaggas	40.148	35.762	40	35.817	17
			(167.04)	(115.323)	
Hanaknak	40	35.817	40.15	35.761	17
			(0.25)	(0.522)	
Hurama	38.261	37.114	39.139	38.226	138
			(1.221)	(4.488)	
Malitta	39.363	33.787	38.888	35.282	141
			(0.515)	(0.141)	
Mamma	37.583	36.933	38.02	36.503	61
			(1.37)	(0.782)	
Salatuwar	39.655	31.994	39.561	33.356	120
			(0.959)	(0.604)	
Samuha	39.619	36.528	38.3	37.118	155
			(0.09)	(0.563)	
Timelkiya	38.027	38.234	38.261	37.114	102
			(0.019)	(0.045)	
Ulama	38.411	33.834	39.835	33.234	167
			(0.4)	(1.277)	
Unipsum	38.021	36.503	37.583	36.933	61
			(389261.721)	(646044.057)	
Wahsusana	39.584	33.418	39.003	31.926	146
			(10.185)	(50.561)	
Zimishuna	40.461	35.65	39.234	34.213	186
			(23.126)	(50.515)	
Mean					116
Median					136

Notes: This table presents the results from a robustness check of our 'proof-of-concept' exercise in Appendix Table 1. For each line, we set the distance elasticity at $\zeta = 1.9$, and using (8), we estimate the geo-coordinates of one fictitiously lost city alongside the ten truly lost cities. We also re-estimate all other parameters of the model. For *Karahna* and *Zimišhuna*, our minimization algorithm hits the non-negativity constraint on their α's. For all known cities, we list the true geo-coordinates of each city, their estimated geo-coordinates from estimating (8), and the distance, in kms, between the true and estimated locations. All latitudes are North, and all longitudes are East. Robust (White) standard errors in parentheses.

Appendix Table 3: **Assigning Lost Cities to Archaeological Sites**

Lost city gravity estimate	Candidate site	Distance to gravity estimate (in km)	Log(p.d.f.)
Durhumit 40.47, 35.65			
	Ayvalıpınar 40.46, 35.65	0.97	1.68
	Oluz Höyük 40.55, 35.63	8.62	-45.66
	Ferzant 40.6, 35.38	27.53	-114.33
	Doğantepe 40.6, 35.6	14.71	-130.09
	Boyalı 40.31, 34.26	123.58	-463.26
Hahhum 38.43, 38.04			
	Imikuşağı 38.52, 38.46	37.9	0.32
	Değirmentepe 38.48, 38.45	36.18	0.26
	Imamoğlu 38.48, 38.48	38.92	0.18
	Arslantepe 38.38, 38.36	28.69	0.17
	Yassıhöyük (Tanır) 38.39, 36.91	98.88	-3.38
Kuburnat 40.71, 36.52			
	Tekkeköy (Samsun) 41.2, 36.45	54.57	-1.01
	Dündartepe 41.25, 36.35	61.74	-1.12
	Kaledoruğu (Kavak) 41.08, 36.04	58.61	-1.26
	Kayapınar Höyüğü 40.16, 36.25	65.42	-1.27
	Bolus (Aktepe) 40.07, 36.5	71.72	-1.28

Continued on the next page

Appendix Table 3: **Assigning Lost Cities to Archaeological Sites (continued)**

Lost city gravity estimate	Candidate site	Distance to gravity estimate (in km)	Log(p.d.f.)
Ninassa 38.98, 34.61			
	Suluca Karahöyük (Hacibektaş) 38.93, 34.55	7.42	0.11
	Topakhöyük 38.61, 34.29	49.37	-0.12
	Zank 38.95, 34.79	16.05	-0.19
	Topaklı 39.01, 34.83	18.93	-0.2
	Uşaklı/Kuşaklı Höyük 39.8, 35.1	100.87	-0.45
Purushaddum 39.71, 32.87			
	Karaoğlan 39.73, 32.83	4.3	-1.13
	Külhöyük (Haymana) 39.48, 32.67	31.43	-1.18
	Ballıkuyumcu 39.77, 32.52	31.77	-1.51
	Çomaklı / İlmez 37.72, 32.5	223.91	-2.21
	Ortakaraviran II 37.38, 32.09	267.98	-2.27
Sinahuttum 39.96, 34.87			
	Yassıhöyük (Yozgat) 39.99, 34.88	3.97	2.05
	Çengeltepe 39.84, 34.87	12.99	1.79
	Eskiyapar 40.16, 34.77	23.91	-0.44
	Mercimektepe 40.88, 35.34	110.6	-2.25
	Suluca Karahöyük (Hacibektaş) 38.93, 34.55	116.96	-3.73

Continued on the next page

Appendix Table 3: **Assigning Lost Cities to Archaeological Sites (continued)**

Lost city gravity estimate	Candidate site	Distance to gravity estimate (in km)	Log(p.d.f.)
Suppiluliya 40.02, 34.62			
	Suluca Karahöyük (Hacibektaş) 38.93, 34.55	120.98	-3.65
	Alaca Höyük 40.23, 34.68	24.3	-51.97
	Büyüknefes (Bronze Age Site) 39.85, 34.5	21.61	-201.29
	Eskiyapar 40.16, 34.77	20.46	-373.9
	Arslantepe 38.38, 38.36	375.37	-Inf
Tuhpiya 39.61, 35.2			
	Çadır (Sorgun) 39.68, 35.14	9.27	1.25
	Uşaklı/Kuşaklı Höyük 39.8, 35.1	22.76	0.34
	Çengeltepe 39.84, 34.87	38.43	-1.48
	Boğazlıyan / Yoğunhisar 39.17, 35.23	49.42	-3.01
	Üyük 40.15, 35.85	82.78	-3.27
Washaniya 39.16, 34.31			
	Yassıhöyük (Çoğun / Kırşehir) 39.32, 34.08	27.42	0.34
	Suluca Karahöyük (Hacibektaş) 38.93, 34.55	32.51	0.08
	Harmandalı 38.95, 33.95	39.14	-0.38
	Zank 38.95, 34.79	48.08	-1.08
	Topaklı 39.01, 34.83	48.17	-1.11

Continued on the next page

Appendix Table 3: **Assigning Lost Cities to Archaeological Sites (continued)**

Lost city gravity estimate	Candidate site	Distance to gravity estimate (in km)	Log(p.d.f.)
Zalpa 38.81, 37.86			
	Yassıhöyük (Tanır) 38.39, 36.91	95.1	-0.86
	Yalak (Boz Höyük) 38.3, 36.44	137.32	-1.39
	Sarız 38.47, 36.5	125.22	-1.67
	Imikuşağı 38.52, 38.46	60.94	-2.32
	Değirmentepe 38.48, 38.45	62.95	-2.56

Notes: This table gives a list of the five most likely potential archaeological sites for each lost city. The table lists each lost city, with its the geo-coordinates (first latitude, North, then longitude, East) derived from estimating our gravity model (for instance, the estimated coordinates for *Durhumit* are 40.47 degrees North and 35.65 degrees East). It then lists the top five candidate sites, with their corresponding geo-coordinates. The sites are ordered in decreasing order of probability density, where we evaluate our estimated probability density, $\hat{f}_l(\varphi, \lambda)$ for each lost city l at each potential site's coordinates, using equation (13). For each potential site, we display the distance (in km) from its corresponding lost city gravity estimate, and the (log) probability density for that site. For instance, Ayvalıpınar is the most likely location for *Durhumit*, at a distance of 0.97km from our gravity estimate, and with a log(p.d.f.) of 1.68.

Appendix Table 4: **Determinants of Ancient City Sizes, Robustness**

	$\log\left(PopT^{1/\theta}\vert_{ancient}\right)$				
	(1)	(2)	(3)	(4)	(5)
Panel A: Barjamovic (2011) locations					
$\log(NaturalRoads)$	0.141			0.163	
	(0.281)			(0.234)	
$\log(RomanRoads)$		0.776*			0.783*
		(0.068)			(0.076)
$\log(Ruggedness)$			0.067	0.135	0.080
			(0.766)	(0.557)	(0.679)
N	25	25	25	25	25
R^2	0.062	0.145	0.006	0.084	0.153
Panel B: known cities only					
$\log(NaturalRoads)$	2.105*			2.526*	
	(0.086)			(0.057)	
$\log(RomanRoads)$		4.540			4.771
		(0.207)			(0.125)
$\log(Ruggedness)$			2.722	3.970***	2.878*
			(0.123)	(0.001)	(0.059)
N	15	15	15	15	15
R^2	0.311	0.126	0.111	0.536	0.251

Notes: This table replicates the results in table IV. Panel A uses the locations of lost cities proposed by Barjamovic (2011) instead of our structural gravity estimates. Panel B uses only the subsample of cities with known locations. Robust p-values are in parentheses.

Appendix Table 5: **Persistence of Economic Activity across 4000 Years, Robustness**

	log(*Population*)			log(*NightLights*)		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Barjamovic (2011) locations						
$\log\left(PopT^{1/\theta}\vert_{ancient}\right)$	0.691		0.899	0.675**		0.889***
	(0.194)		(0.143)	(0.014)		(0.002)
$\log\left(CropYield\right)$		0.518	1.123		0.558	1.158*
		(0.681)	(0.390)		(0.408)	(0.086)
N	25	25	25	25	25	25
R^2	0.045	0.007	0.073	0.134	0.025	0.229
Panel B: known cities only						
$\log\left(PopT^{1/\theta}\vert_{ancient}\right)$	0.171*		0.209**	0.066		0.091
	(0.062)		(0.026)	(0.203)		(0.120)
$\log\left(CropYield\right)$		1.138	1.866		0.935	1.253*
		(0.458)	(0.169)		(0.254)	(0.088)
N	15	15	15	15	15	15
R^2	0.135	0.035	0.223	0.062	0.073	0.184
Panel C: ancient site matched to the largest modern settlement						
$\log\left(PopT^{1/\theta}\vert_{ancient}\right)$	0.235**		0.303**			
	(0.031)		(0.013)			
$\log\left(CropYield\right)$		0.716	1.791*			
		(0.516)	(0.078)			
N	24	24	24			
R^2	0.154	0.015	0.236			
Panel D: ancient site matched to the closest modern settlement						
$\log\left(PopT^{1/\theta}\vert_{ancient}\right)$	0.252**		0.311***			
	(0.019)		(0.010)			
$\log\left(CropYield\right)$		0.478	1.582			
		(0.648)	(0.112)			
N	24	24	24			
R^2	0.207	0.008	0.283			

Notes: This table replicates the results in table V. Panel A uses the locations of lost cities proposed by Barjamovic (2011) instead of our structural gravity estimates. Since none of these locations correspond to the modern city of Ankara, this estimation features the full set of 25 cities. See text for details. Panel B uses only the subsample of 15 cities with known locations, both in estimation and in backing out $PopT^{1/\theta}\vert_{ancient}$ values. Instead of matching ancient settlements to the *sum of all* modern settlements within 20 km, Panel C and D match them to the largest modern urban settlement within 20 km and to the closest one, respectively. Since the dependent variables in Panels C and D follow administrative boundaries, the *NighLights* variable does not apply to these specifications. Robust p-values are in parentheses.

Appendix Table 6: **Structural versus Naive Gravity: Locating Lost Cities**

	(1)	(2)	(3)
Durhumit	127	48	99
Hahhum	193	102	256
Kuburnat	62	70	70
Ninassa	81	93	149
Purushaddum	241	193	432
Sinahuttum	34	24	35
Suppiluliya	67	85	37
Tuhpiya	200	112	99
Washaniya	132	13	143
Zalpa	96	131	224
Mean	123.3	87.1	154.4

Notes: This table compares the estimated locations of lost cities using either our structural gravity model (8) or a naive gravity model (17) similar to that used by Tobler and Wineburg (1971). Column 1 reports the distance between the structural and naive estimates for lost city locations. Column 2 gives the distance between the location proposed by Barjamovic (2011) and the structural estimates. Column 3 gives the distance between the location proposed by Barjamovic (2011) and the naive estimates. All distances are in km.

Appendix Table 7: **Structural versus Naive Gravity: Ancient City Sizes**

	Structural estimates				Naive estimates			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Determinants of ancient city sizes								
				$\log\left(PopT^{1/\theta}\vert_{ancient}\right)$				
$\log(NaturalRoads)$	1.404**	1.783***	1.163*	1.505**	0.630***	0.678***	0.682	1.097*
	(0.013)	(0.002)	(0.092)	(0.038)	(0.000)	(0.000)	(0.301)	(0.053)
$\log(Ruggedness)$		3.189***		2.147**		0.449		1.022**
		(0.000)		(0.012)		(0.145)		(0.025)
N	25	25	10	10	25	25	10	10
R^2	0.224	0.508	0.178	0.378	0.296	0.348	0.124	0.508
Sample of cities	all	all	lost	lost	all	all	lost	lost
Panel B: Persistence of economic activity across 4000 years								
				$\log(Population)$				
$\log\left(PopT^{1/\theta}\vert_{ancient}\right)$	0.230**	0.297**	0.387**	0.533**	0.313	0.307	-0.139	-0.140
	(0.035)	(0.015)	(0.035)	(0.063)	(0.376)	(0.367)	(0.717)	(0.677)
$\log(CropYield)$		1.781*		2.238*		0.921		-0.008
		(0.079)		(0.093)		(0.439)		(0.997)
N	24	24	10	10	25	25	10	10
R^2	0.145	0.226	0.362	0.487	0.035	0.059	0.010	0.010
Sample of cities	all	all	lost	lost	all	all	lost	lost

Notes: This table compares the estimated sizes of ancient cities using either our structural gravity model (8) and (11) or a naive gravity model (17) similar to that used by Tobler and Wineburg (1971). Panel A replicates the results in table IV, using either our structural estimates (columns 1-4) or naive estimates (column 5-8). Columns 1 and 2 simply reproduce columns 1 and 4 of table IV for comparison. Columns 3 and 4 replicate the specifications of columns 1 and 2 on the subset of lost cities only. Column 5 to 8 replicate the specifications of columns 1 to 4 using naive estimates instead of structural ones. Panel B replicates the results in table V, using either our structural estimates (columns 1-4) or naive estimates (column 5-8). Columns 1 and 2 simply reproduce columns 1 and 3 of table V for comparison. Columns 3 and 4 replicate the specifications of columns 1 and 2 on the subset of lost cities only. Column 5 to 8 replicate the specifications of columns 1 to 4 using naive estimates instead of structural ones. To offer a meaningful comparison between structural and naive estimates, we do not drop *Purušhaddum* in columns 3-4 and 7-8, as it is an outlier (near modern Ankara) for the structural estimates (columns 3-4), but not for the naive ones (columns 7-8). Robust p-values are in parentheses.