The Out of Africa Hypothesis, Genetic Diversity and Comparative Development

Quamrul Ashraf and Oded Galor

October 17, 2012

Unified Growth Theory

Main Hypothesis

- Deep-rooted factors, determined tens of thousands of years ago, have had a significant effect on the:
  - process of development
  - contemporary differences in income per capita across countries
- Variation in migratory distance from East Africa to indigenous settlements across the globe
  - $\Longrightarrow$ distribution of genetic diversity $\Longrightarrow$ overall diversity $\Longrightarrow$ comparative development

Main Hypothesis

- Deep-rooted factors, determined tens of thousands of years ago, have had a significant effect on the:
  - process of development
  - contemporary differences in income per capita across countries
- Variation in migratory distance from East Africa to indigenous settlements across the globe
  - $\Longrightarrow$ distribution of genetic diversity $\Longrightarrow$ overall diversity $\Longrightarrow$ comparative development

Main Hypothesis

- Deep-rooted factors, determined tens of thousands of years ago, have had a significant effect on the:
  - process of development
  - contemporary differences in income per capita across countries
- Variation in migratory distance from East Africa to indigenous settlements across the globe
  - $\Longrightarrow$ distribution of genetic diversity $\Longrightarrow$ overall diversity $\Longrightarrow$ comparative development

Main Hypothesis

- Deep-rooted factors, determined tens of thousands of years ago, have had a significant effect on the:
  - process of development
  - contemporary differences in income per capita across countries
- Variation in migratory distance from East Africa to indigenous settlements across the globe
  - $\Longrightarrow$ distribution of genetic diversity $\Longrightarrow$ overall diversity $\Longrightarrow$ comparative development

Main Hypothesis

- Deep-rooted factors, determined tens of thousands of years ago, have had a significant effect on the:
  - process of development
  - contemporary differences in income per capita across countries
- Variation in migratory distance from East Africa to indigenous settlements across the globe
  - $\Longrightarrow$ distribution of genetic diversity $\Longrightarrow$ overall diversity $\Longrightarrow$ comparative development

Main Hypothesis

- Deep-rooted factors, determined tens of thousands of years ago, have had a significant effect on the:
  - process of development
  - contemporary differences in income per capita across countries
- Variation in migratory distance from East Africa to indigenous settlements across the globe
  - $\Longrightarrow$ distribution of genetic diversity $\Longrightarrow$ overall diversity $\Longrightarrow$ comparative development

Main Hypothesis

- Deep-rooted factors, determined tens of thousands of years ago, have had a significant effect on the:
  - process of development
  - contemporary differences in income per capita across countries
- Variation in migratory distance from East Africa to indigenous settlements across the globe
  - $\Longrightarrow$ distribution of genetic diversity $\Longrightarrow$ overall diversity $\Longrightarrow$ comparative development

Main Building Blocks of the Proposed Hypothesis

- The Serial-Founder Effect: Lower genetic diversity exists among indigenous populations at greater migratory distances from east Africa
- Trade-offs associated with diversity within a society $\Longrightarrow$ optimal level of diversity (may differ in different stages of development)

Main Building Blocks of the Proposed Hypothesis

- The Serial-Founder Effect: Lower genetic diversity exists among indigenous populations at greater migratory distances from east Africa
- Trade-offs associated with diversity within a society $\Longrightarrow$ optimal level of diversity (may differ in different stages of development)

The Out of Africa Hypothesis and Genetic Diversity

- The Serial-Founder Effect:
  - In the course of human expansion out of Africa, as subgroups of the populations of parental colonies left to establish new settlements further away, they carried with them only a subset of the overall genetic diversity

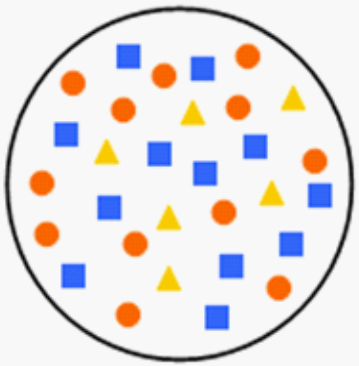
$\Longrightarrow$ Lower genetic diversity exists among indigenous populations at greater migratory distances from east Africa

The Out of Africa Hypothesis and Genetic Diversity

- The Serial-Founder Effect:
  - In the course of human expansion out of Africa, as subgroups of the populations of parental colonies left to establish new settlements further away, they carried with them only a subset of the overall genetic diversity

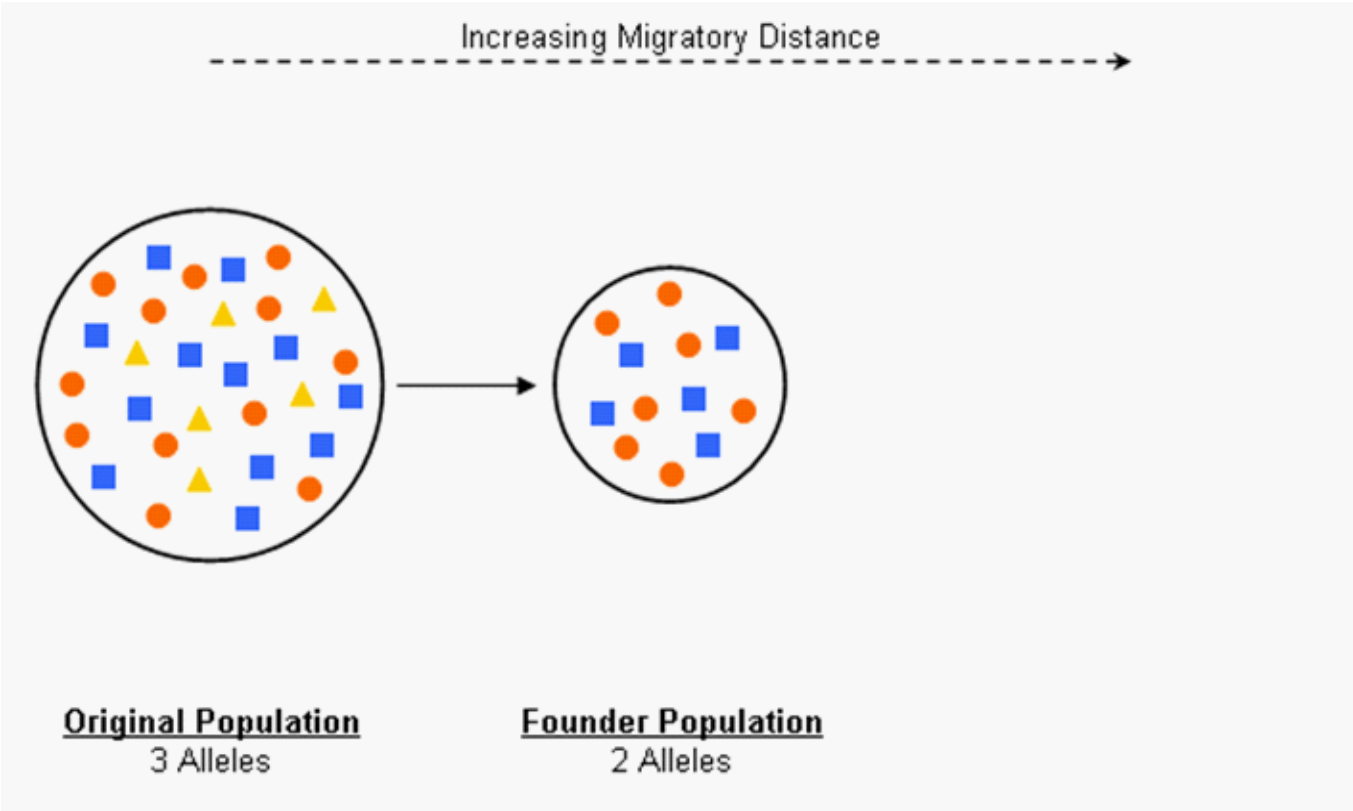
$\Longrightarrow$ Lower genetic diversity exists among indigenous populations at greater migratory distances from east Africa

An Illustration of the Serial-Founder Effect

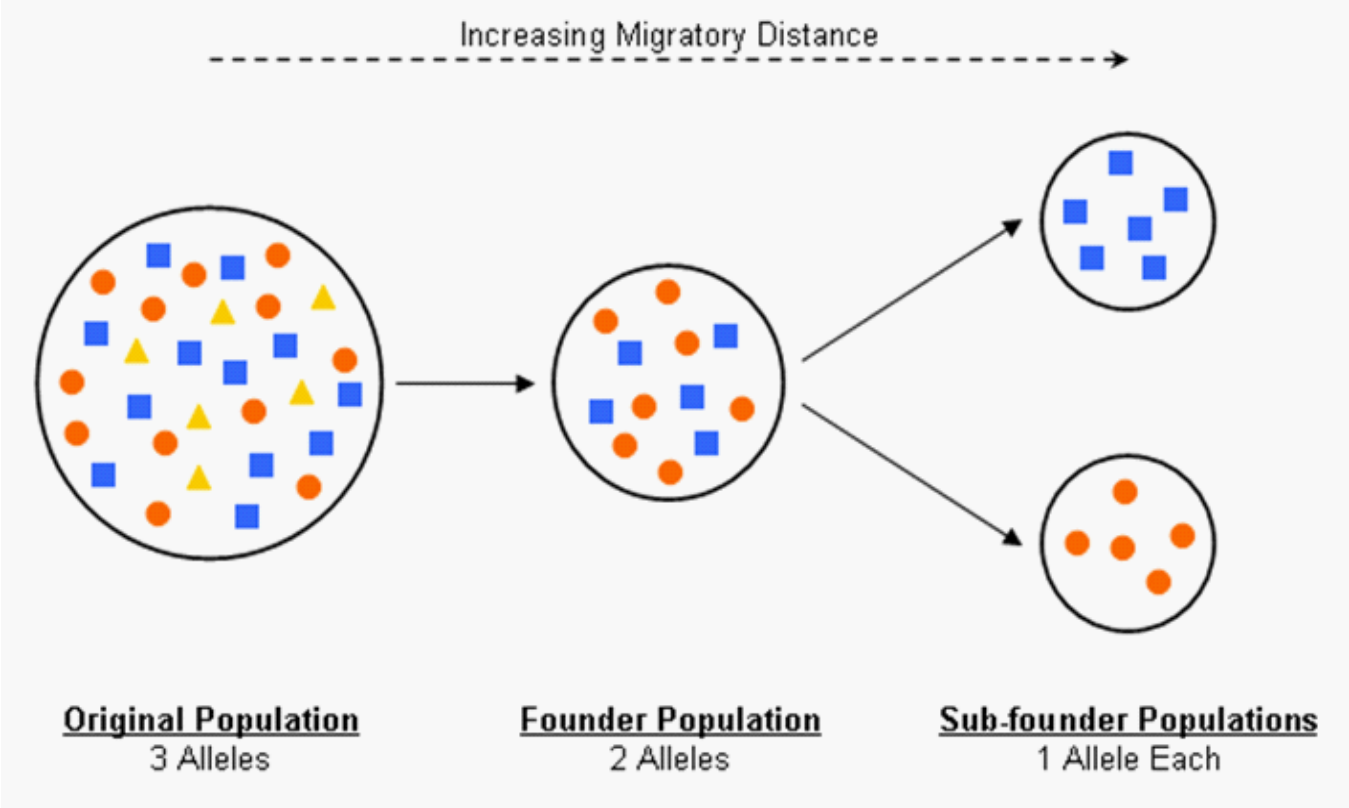


Original Population
3 Alleles

An Illustration of the Serial-Founder Effect



An Illustration of the Serial-Founder Effect



Measurement of Genetic Diversity

- Population geneticists measure genetic diversity using an index called "expected heterozygosity"
- The index captures the probability that two individuals, selected at random from the relevant population, are genetically different from one another
- Construction of Expected Heterozygosity:
  - Gene-specific Herfindahl Index reflecting its allelic ("gene variant") frequencies
  - Averaged over large number of genes

Measurement of Genetic Diversity

- Population geneticists measure genetic diversity using an index called "expected heterozygosity"
- The index captures the probability that two individuals, selected at random from the relevant population, are genetically different from one another
- Construction of Expected Heterozygosity:
  - Gene-specific Herfindahl Index reflecting its allelic ("gene variant") frequencies
  - Averaged over large number of genes

Measurement of Genetic Diversity

- Population geneticists measure genetic diversity using an index called "expected heterozygosity"
- The index captures the probability that two individuals, selected at random from the relevant population, are genetically different from one another
- Construction of Expected Heterozygosity:
  - Gene-specific Herfindahl Index reflecting its allelic ("gene variant") frequencies
  - Averaged over large number of genes

Measurement of Genetic Diversity

- Population geneticists measure genetic diversity using an index called "expected heterozygosity"
- The index captures the probability that two individuals, selected at random from the relevant population, are genetically different from one another
- Construction of Expected Heterozygosity:
  - Gene-specific Herfindahl Index reflecting its allelic ("gene variant") frequencies
  - Averaged over large number of genes

Measurement of Genetic Diversity

- Population geneticists measure genetic diversity using an index called "expected heterozygosity"
- The index captures the probability that two individuals, selected at random from the relevant population, are genetically different from one another
- Construction of Expected Heterozygosity:
  - Gene-specific Herfindahl Index reflecting its allelic ("gene variant") frequencies
  - Averaged over large number of genes

The Expected Heterozygosity Index

- For a single gene λ with k_λ alleles in a given population, the locus-specific heterozygosity is

$$H_\lambda = 1 - \sum_{i=1}^{k_\lambda} (p_i^\lambda)^2$$

where p_i^λ is the observed frequency of the i-th allele of locus λ

- Averaging over multiple loci, the expected heterozygosity (genetic diversity) of the population is:

$$H = \frac{1}{m} \sum_{\lambda=1}^{m} H_\lambda = 1 - \frac{1}{m} \sum_{\lambda=1}^{m} \sum_{i=1}^{k_\lambda} (p_i^\lambda)^2$$

where m is the number of different loci being considered

The Expected Heterozygosity Index

- For a single gene λ with k_λ alleles in a given population, the locus-specific heterozygosity is

$$H_\lambda = 1 - \sum_{i=1}^{k_\lambda} (p_i^\lambda)^2$$

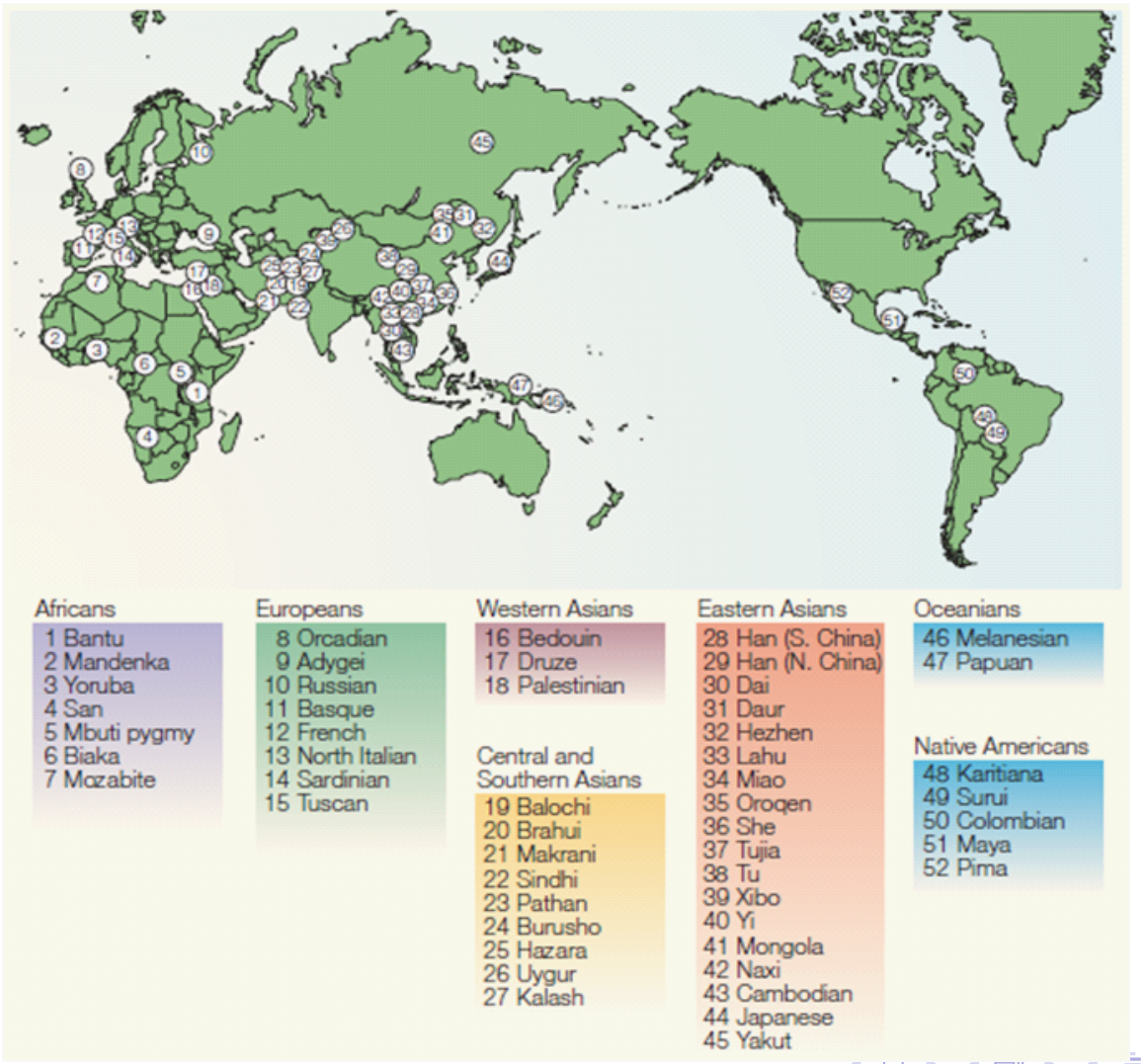
where p_i^λ is the observed frequency of the i-th allele of locus λ

- Averaging over multiple loci, the expected heterozygosity (genetic diversity) of the population is:

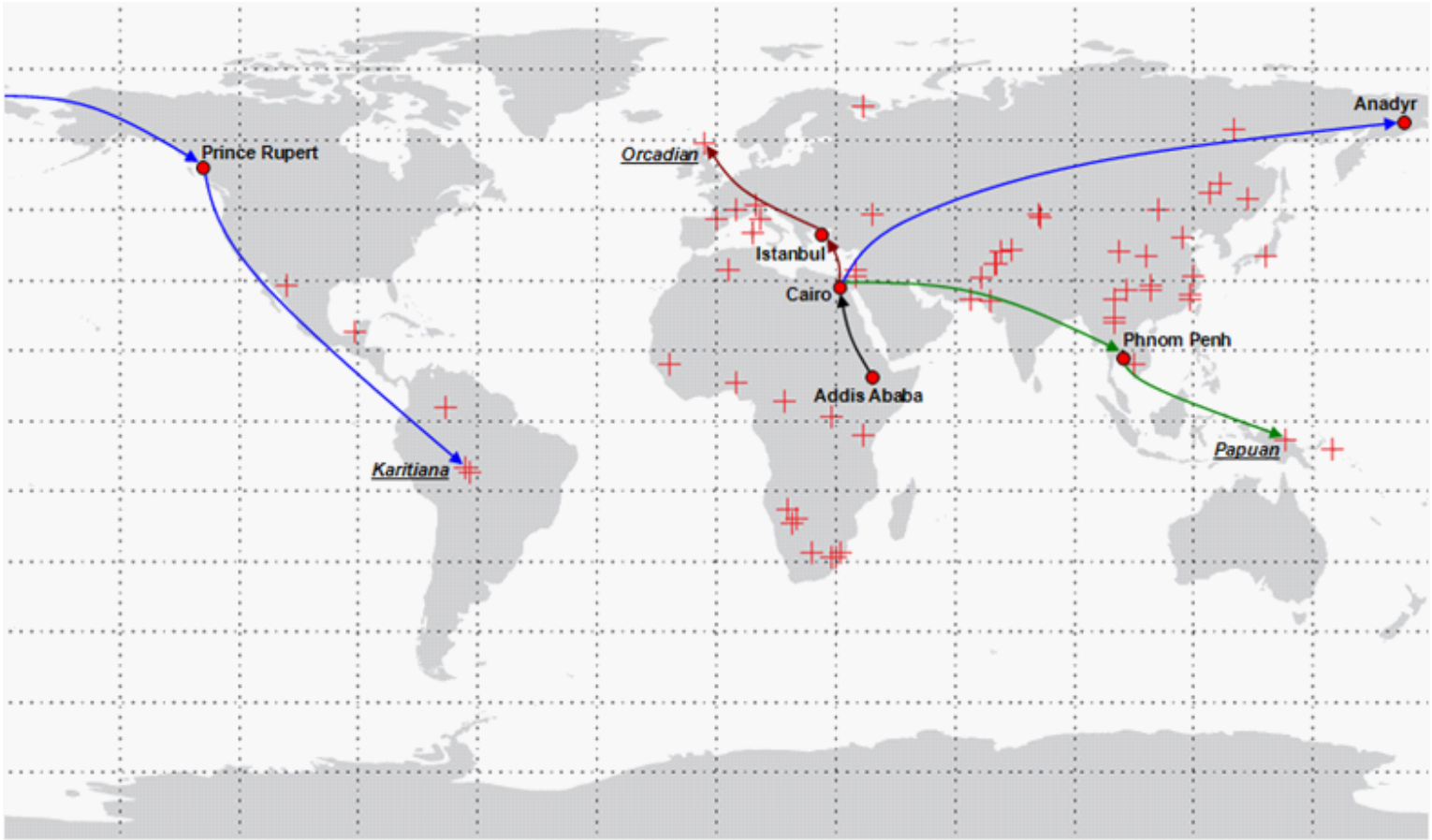
$$H = \frac{1}{m} \sum_{\lambda=1}^{m} H_\lambda = 1 - \frac{1}{m} \sum_{\lambda=1}^{m} \sum_{i=1}^{k_\lambda} (p_i^\lambda)^2$$

where m is the number of different loci being considered

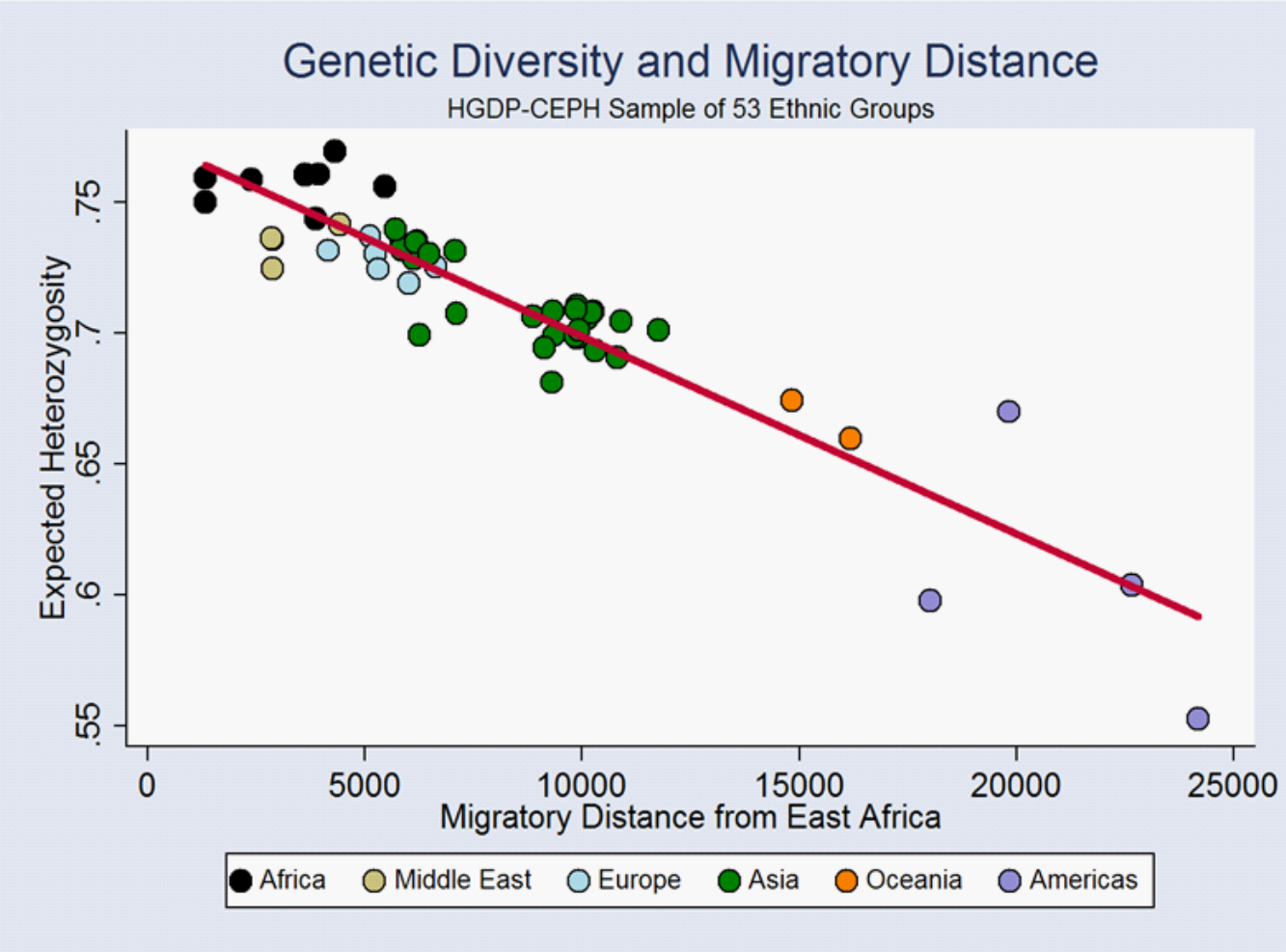
The HGDP Ethnic Groups



Migratory Distance from East Africa



Evidence of the Serial-Founder Effect



Expected Heterozygosity in Microsattelite

- Expected heterozygosity calculated for the 53 ethnic groups in the HGDP using allelic frequencies for 783 microsattelite loci
- Microsattelites – a class of non-protein-coding regions of the human genome:
  - Selectively neutral – ensures that the observed cross-sectional variation in diversity is not due to differential forces of natural selection
  - Mutationally active – facilitates the construction of "population trees" and thus the genealogical and migratory histories of populations

Expected Heterozygosity in Microsattelite

- Expected heterozygosity calculated for the 53 ethnic groups in the HGDP using allelic frequencies for 783 microsattelite loci
- Microsattelites – a class of non-protein-coding regions of the human genome:
  - Selectively neutral – ensures that the observed cross-sectional variation in diversity is not due to differential forces of natural selection
  - Mutationally active – facilitates the construction of "population trees" and thus the genealogical and migratory histories of populations

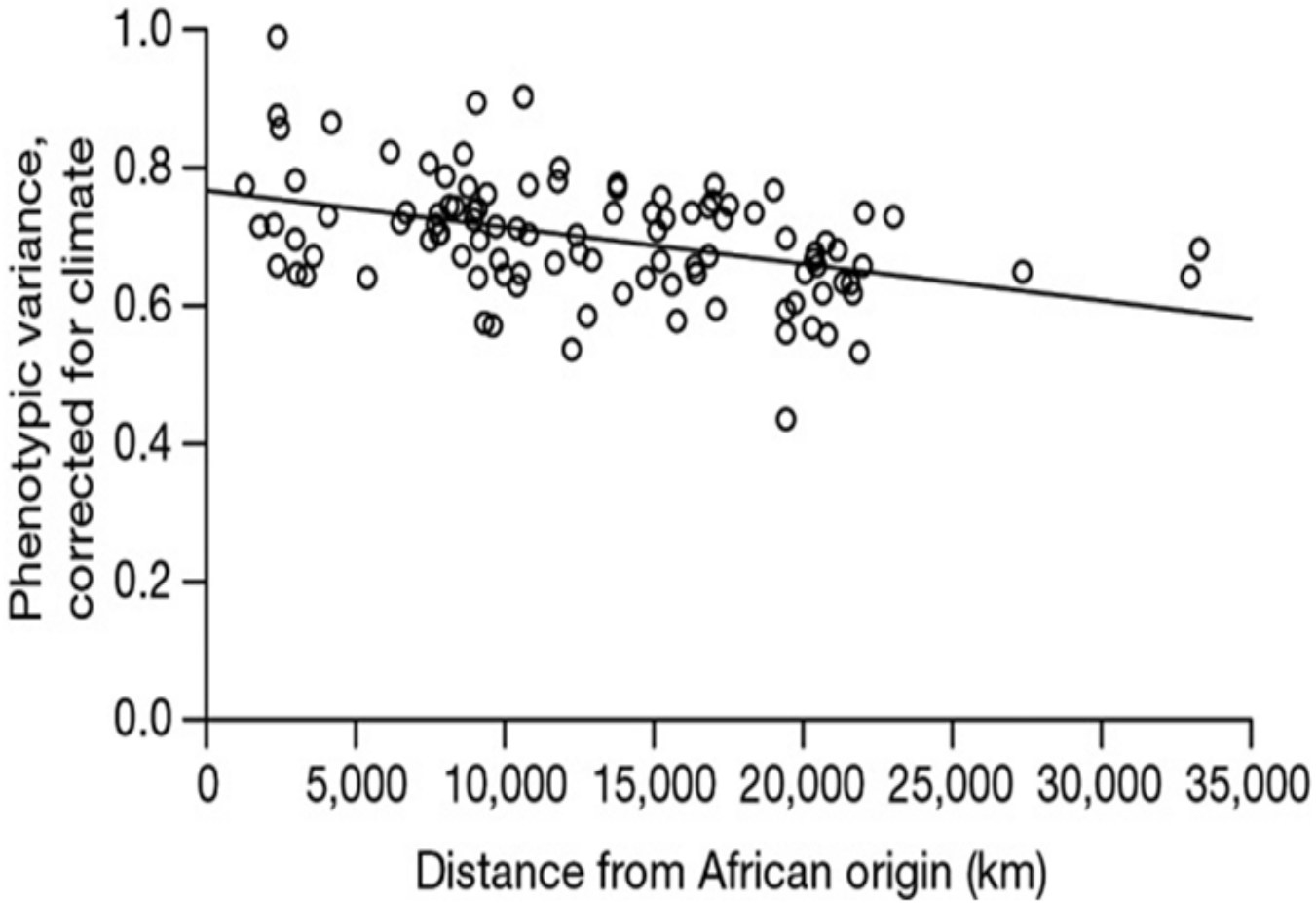
Expected Heterozygosity in Microsattelite

- Expected heterozygosity calculated for the 53 ethnic groups in the HGDP using allelic frequencies for 783 microsattelite loci
- Microsattelites – a class of non-protein-coding regions of the human genome:
  - Selectively neutral – ensures that the observed cross-sectional variation in diversity is not due to differential forces of natural selection
  - Mutationally active – facilitates the construction of "population trees" and thus the genealogical and migratory histories of populations

Expected Heterozygosity in Microsattelite

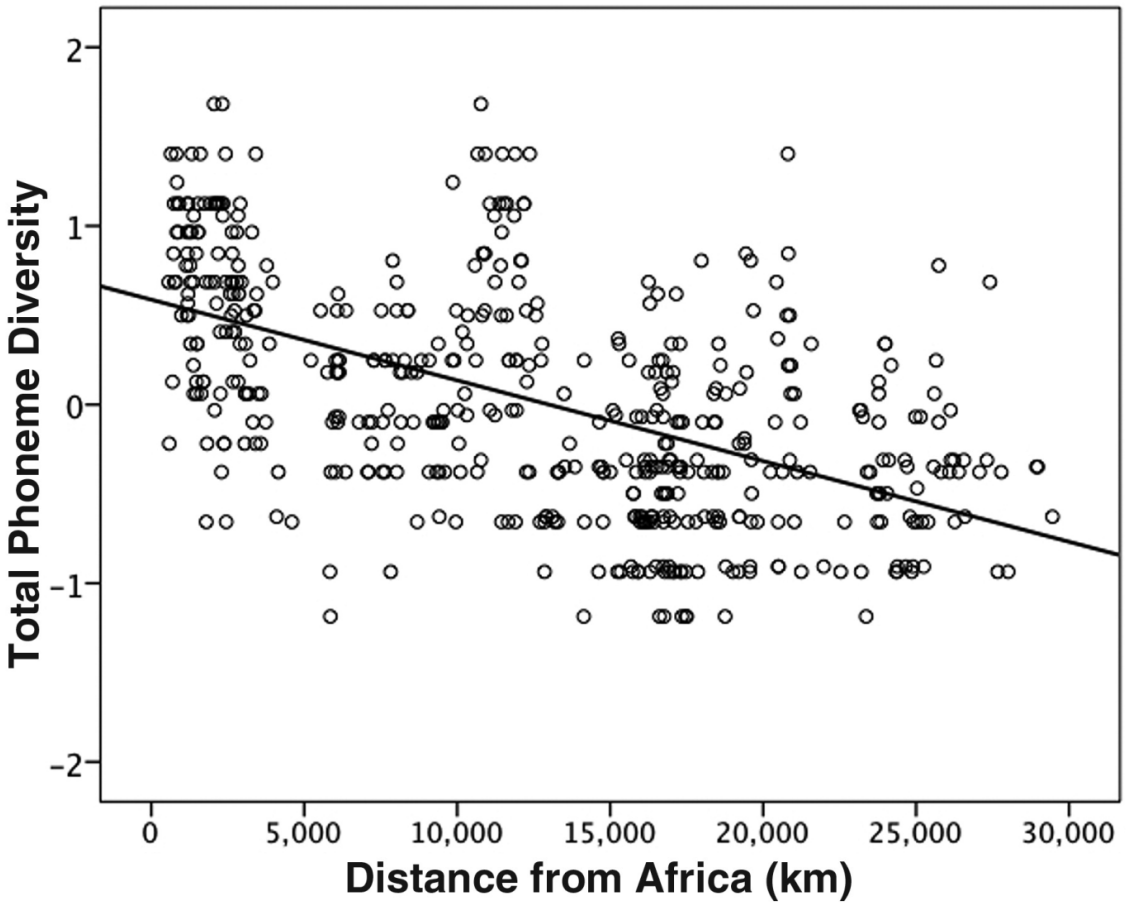
- Expected heterozygosity calculated for the 53 ethnic groups in the HGDP using allelic frequencies for 783 microsattelite loci
- Microsattelites – a class of non-protein-coding regions of the human genome:
  - Selectively neutral – ensures that the observed cross-sectional variation in diversity is not due to differential forces of natural selection
  - Mutationally active – facilitates the construction of "population trees" and thus the genealogical and migratory histories of populations

Positive Relationship: Heterozygosity in Microsattelite and Phenotypic Diversity



Source: (Manica et al. *Nature*, 2007)

Positive Relationship: Heterozygosity in Microsattelite and Cultural Diversity



Source: Atkinson, *Science*, 2011

Diversity and Economic Development

- Diversity within a society generates costs and benefits:
  - Costs: diversity reduces the likelihood of cooperative or trustful behavior

    $\Longrightarrow$ Diversity generates inefficiencies in the operation of the economy relative to its PPF
  - Benefits: wider spectrum of traits is likely to be complementary to the adoption or implementation of new technologies
    - Variety permits specialization in complementary tasks
    - Greater variance of cognitive traits increases innovations in an environment in which only high levels of cognitive abilities contribute to technological innovation

    $\Longrightarrow$ Diversity enhances, productivity, knowledge creation and fosters technological progress

Diversity and Economic Development

- Diversity within a society generates costs and benefits:
  - Costs: diversity reduces the likelihood of cooperative or trustful behavior

$$\Longrightarrow$$ Diversity generates inefficiencies in the operation of the economy relative to its PPF

  - Benefits: wider spectrum of traits is likely to be complementary to the adoption or implementation of new technologies
    - Variety permits specialization in complementary tasks
    - Greater variance of cognitive traits increases innovations in an environment in which only high levels of cognitive abilities contribute to technological innovation

$$\Longrightarrow$$ Diversity enhances, productivity, knowledge creation and fosters technological progress

Diversity and Economic Development

- Diversity within a society generates costs and benefits:
  - Costs: diversity reduces the likelihood of cooperative or trustful behavior

    $\Longrightarrow$ Diversity generates inefficiencies in the operation of the economy relative to its PPF

  - Benefits: wider spectrum of traits is likely to be complementary to the adoption or implementation of new technologies
    - Variety permits specialization in complementary tasks
    - Greater variance of cognitive traits increases innovations in an environment in which only high levels of cognitive abilities contribute to technological innovation

    $\Longrightarrow$ Diversity enhances, productivity, knowledge creation and fosters technological progress

Diversity and Economic Development

- Diversity within a society generates costs and benefits:
  - Costs: diversity reduces the likelihood of cooperative or trustful behavior

    $\Longrightarrow$ Diversity generates inefficiencies in the operation of the economy relative to its PPF

  - Benefits: wider spectrum of traits is likely to be complementary to the adoption or implementation of new technologies
    - Variety permits specialization in complementary tasks
    - Greater variance of cognitive traits increases innovations in an environment in which only high levels of cognitive abilities contribute to technological innovation

    $\Longrightarrow$ Diversity enhances, productivity, knowledge creation and fosters technological progress

Diversity and Economic Development

- Diversity within a society generates costs and benefits:
  - Costs: diversity reduces the likelihood of cooperative or trustful behavior

    $\Longrightarrow$ Diversity generates inefficiencies in the operation of the economy relative to its PPF

  - Benefits: wider spectrum of traits is likely to be complementary to the adoption or implementation of new technologies
    - Variety permits specialization in complementary tasks
    - Greater variance of cognitive traits increases innovations in an environment in which only high levels of cognitive abilities contribute to technological innovation

      $\Longrightarrow$ Diversity enhances, productivity, knowledge creation and fosters technological progress

Diversity and Economic Development

- Diversity within a society generates costs and benefits:
  - Costs: diversity reduces the likelihood of cooperative or trustful behavior

    $\Longrightarrow$ Diversity generates inefficiencies in the operation of the economy relative to its PPF

  - Benefits: wider spectrum of traits is likely to be complementary to the adoption or implementation of new technologies
    - Variety permits specialization in complementary tasks
    - Greater variance of cognitive traits increases innovations in an environment in which only high levels of cognitive abilities contribute to technological innovation

    $\Longrightarrow$ Diversity enhances, productivity, knowledge creation and fosters technological progress

Diversity and Economic Development

- Diversity within a society generates costs and benefits:
  - Costs: diversity reduces the likelihood of cooperative or trustful behavior

    $\Longrightarrow$ Diversity generates inefficiencies in the operation of the economy relative to its PPF

  - Benefits: wider spectrum of traits is likely to be complementary to the adoption or implementation of new technologies
    - Variety permits specialization in complementary tasks
    - Greater variance of cognitive traits increases innovations in an environment in which only high levels of cognitive abilities contribute to technological innovation

    $\Longrightarrow$ Diversity enhances, productivity, knowledge creation and fosters technological progress

Optimal Diversity

- Diminishing marginal benefits of diversity & homogeneity

$\Longrightarrow$ A hump-shaped relationship across populations between diversity and development outcomes

$\Longrightarrow$ Higher optimal level of diversity in the post-industrialization environment of rapid technological change

Optimal Diversity

- Diminishing marginal benefits of diversity & homogeneity

$\Longrightarrow$ A hump-shaped relationship across populations between diversity and development outcomes

$\Longrightarrow$ Higher optimal level of diversity in the post-industrialization environment of rapid technological change

Optimal Diversity

- Diminishing marginal benefits of diversity & homogeneity

  $\Longrightarrow$ A hump-shaped relationship across populations between diversity and development outcomes

  $\Longrightarrow$ Higher optimal level of diversity in the post-industrialization environment of rapid technological change

Theoretical Foundations of the Hump-Shaped Effect of Diversity

- $\omega \in [0,1] \equiv$ degree of diversity;
- $\omega = 0 \equiv$ homogenous society
- $z \equiv$ institutional, geographical, and human capital factors
- Output per capita

$$y = (1 - \alpha\omega)A(z,\omega)f(x) \equiv y(\omega); \qquad \alpha \in (0,1)$$

Theoretical Foundations of the Hump-Shaped Effect of Diversity

- $\omega \in [0,1] \equiv$ degree of diversity;
- $\omega = 0 \equiv$ homogenous society
- $z \equiv$ institutional, geographical, and human capital factors
- Output per capita

$$y = (1 - \alpha\omega)A(z,\omega)f(x) \equiv y(\omega); \qquad \alpha \in (0,1)$$

Theoretical Foundations of the Hump-Shaped Effect of Diversity

- $\omega \in [0,1] \equiv$ degree of diversity;
- $\omega = 0 \equiv$ homogenous society
- $z \equiv$ institutional, geographical, and human capital factors
- Output per capita

$$y = (1 - \alpha\omega)A(z,\omega)f(x) \equiv y(\omega); \qquad \alpha \in (0,1)$$

Theoretical Foundations of the Hump-Shaped Effect of Diversity

- $\omega \in [0,1] \equiv$ degree of diversity;
- $\omega = 0 \equiv$ homogenous society
- $z \equiv$ institutional, geographical, and human capital factors
- Output per capita

$$y = (1 - \alpha\omega)A(z,\omega)f(x) \equiv y(\omega); \qquad \alpha \in (0,1)$$

Theoretical Foundations of the Hump-Shaped Effect of Diversity

- Diversity and TFP growth

$$A(z,\omega) > 0, \;\; A_\omega(z,\omega) > 0, \quad A_{\omega\omega}(z,\omega) < 0$$

$$\lim_{\omega \longrightarrow 0} A_\omega(z,\omega) = \infty; \;\; \lim_{\omega \longrightarrow 1} A_\omega(z,\omega) = 0$$

e.g.,

$$A(z,\omega) = z \int_0^\omega \omega_i^\theta \, di \qquad \theta \in (0,1)$$

Theoretical Foundations of the Hump-Shaped Effect of Diversity

- Diversity and TFP growth

$$A(z,\omega) > 0, \;\; A_\omega(z,\omega) > 0, \quad A_{\omega\omega}(z,\omega) < 0$$

$$\lim_{\omega \longrightarrow 0} A_\omega(z,\omega) = \infty; \;\; \lim_{\omega \longrightarrow 1} A_\omega(z,\omega) = 0$$

e.g.,

$$A(z,\omega) = z \int_0^\omega \omega_i^\theta \, di \qquad \theta \in (0,1)$$

Theoretical Foundations of the Hump-Shaped Effect of Diversity

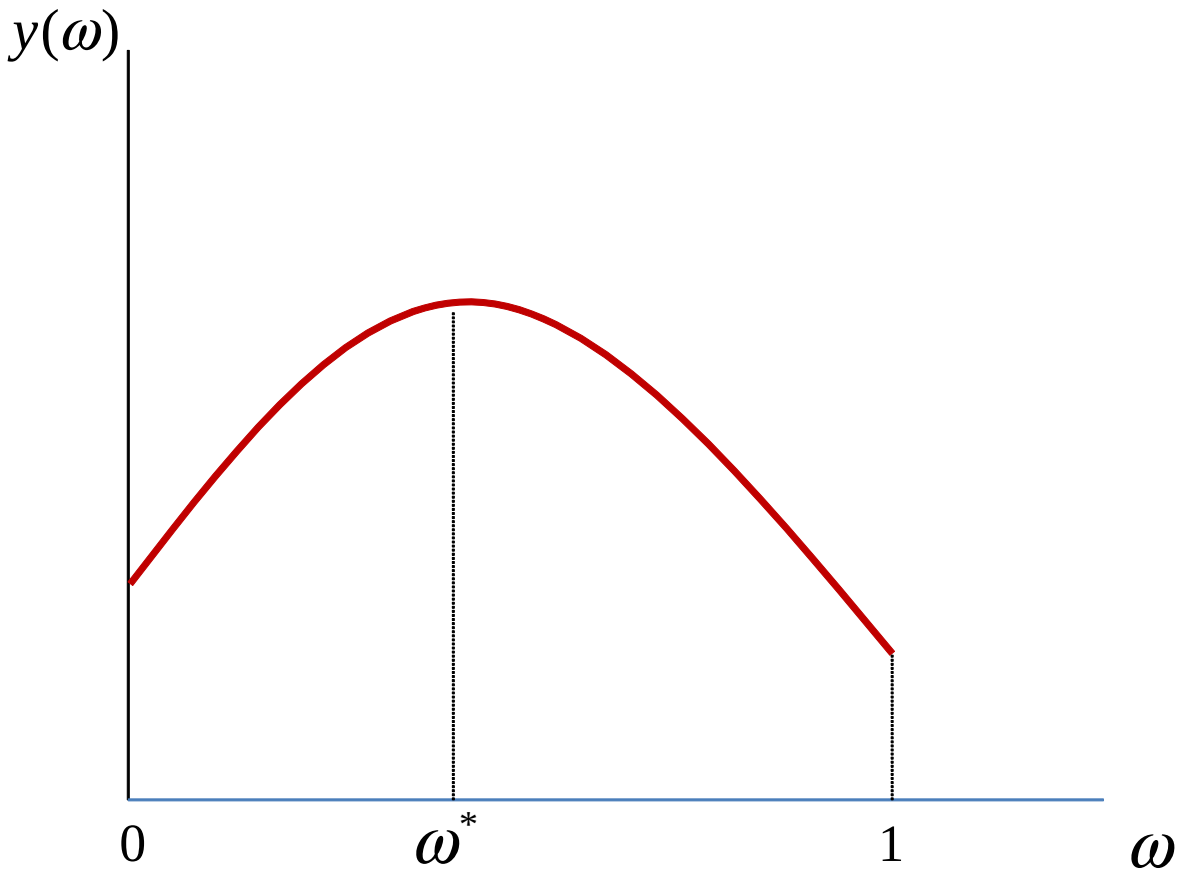
- Properties of $y(\omega)$

$$y'(\omega) = [(1-\alpha\omega)A_{\omega}(z,\omega) - \alpha A(z,\omega)]f(x)$$

$$y''(\omega) = [(1-\alpha\omega)A_{\omega\omega}(z,\omega) - 2\alpha A_{\omega}(z,\omega)]f(x) < 0$$

$$\lim_{\omega \longrightarrow 0} y'(\omega) > 0; \quad \lim_{\omega \longrightarrow 1} y'(\omega) < 0$$

Theoretical Foundations of the Hump-Shaped Effect of Diversity



Historical and Contemporary Analysis

- Stage I (historical analysis):
  - Restrict attention to the pre-colonial era:
    - (a) using actual genetic diversity
    - (b) using projected diversity
- Stage II (contemporary analysis):
  - Accounting for:
    - genetic diversity within each ethnic group of each country
    - genetic distance between ethnic groups of each country

Historical and Contemporary Analysis

- Stage I (historical analysis):
  - Restrict attention to the pre-colonial era:
    - (a) using actual genetic diversity
    - (b) using projected diversity
- Stage II (contemporary analysis):
  - Accounting for:
    - genetic diversity within each ethnic group of each country
    - genetic distance between ethnic groups of each country

Historical and Contemporary Analysis

- Stage I (historical analysis):
  - Restrict attention to the pre-colonial era:
    - (a) using actual genetic diversity
    - (b) using projected diversity
- Stage II (contemporary analysis):
  - Accounting for:
    - genetic diversity within each ethnic group of each country
    - genetic distance between ethnic groups of each country

Historical and Contemporary Analysis

- Stage I (historical analysis):
  - Restrict attention to the pre-colonial era:
    - (a) using actual genetic diversity
    - (b) using projected diversity
- Stage II (contemporary analysis):
  - Accounting for:
    - genetic diversity within each ethnic group of each country
    - genetic distance between ethnic groups of each country

Historical and Contemporary Analysis

- Stage I (historical analysis):
  - Restrict attention to the pre-colonial era:
    - (a) using actual genetic diversity
    - (b) using projected diversity
- Stage II (contemporary analysis):
  - Accounting for:
    - genetic diversity within each ethnic group of each country
    - genetic distance between ethnic groups of each country

Historical and Contemporary Analysis

- Stage I (historical analysis):
  - Restrict attention to the pre-colonial era:
    - (a) using actual genetic diversity
    - (b) using projected diversity
- Stage II (contemporary analysis):
  - Accounting for:
    - genetic diversity within each ethnic group of each country
    - genetic distance between ethnic groups of each country

Historical and Contemporary Analysis

- Stage I (historical analysis):
  - Restrict attention to the pre-colonial era:
    - (a) using actual genetic diversity
    - (b) using projected diversity
- Stage II (contemporary analysis):
  - Accounting for:
    - genetic diversity within each ethnic group of each country
    - genetic distance between ethnic groups of each country

Comparative Development in the Pre-Colonial Era

- The relevant outcome for comparative development is population density
- Population density should, in part, be affected by:
  - Land productivity – The Geographic channel
  - The timing of the Neolithic Revolution - The Diamond channel

Comparative Development in the Pre-Colonial Era

- The relevant outcome for comparative development is population density
- Population density should, in part, be affected by:
  - Land productivity – The Geographic channel
  - The timing of the Neolithic Revolution - The Diamond channel

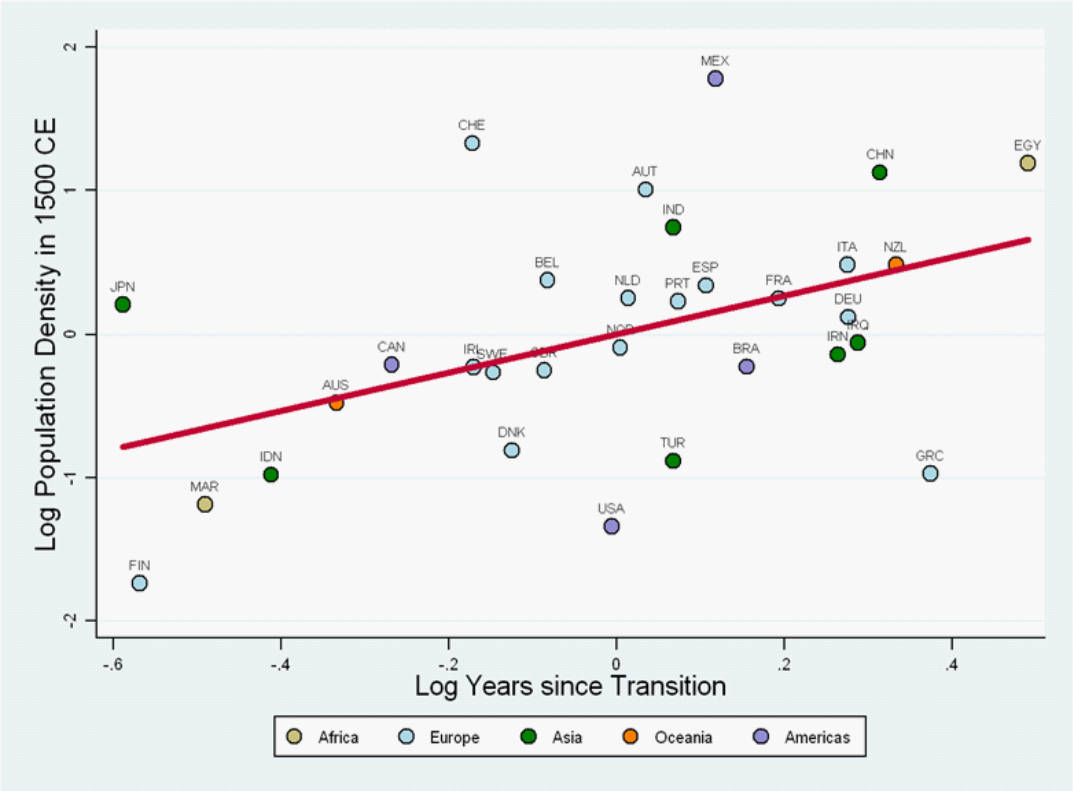
Comparative Development in the Pre-Colonial Era

- The relevant outcome for comparative development is population density
- Population density should, in part, be affected by:
  - Land productivity – The Geographic channel
  - The timing of the Neolithic Revolution - The Diamond channel

Comparative Development in the Pre-Colonial Era

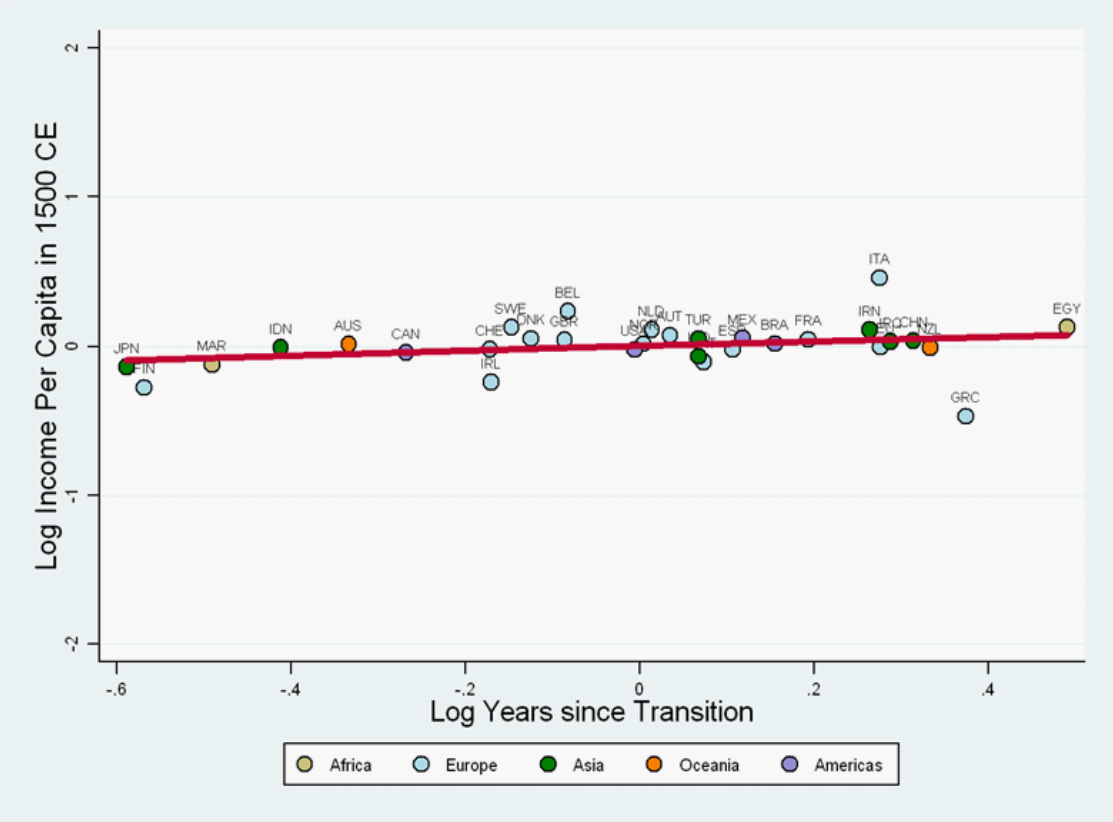
- The relevant outcome for comparative development is population density
- Population density should, in part, be affected by:
  - Land productivity – The Geographic channel
  - The timing of the Neolithic Revolution - The Diamond channel

Transition Timing and Population Density in 1500 CE



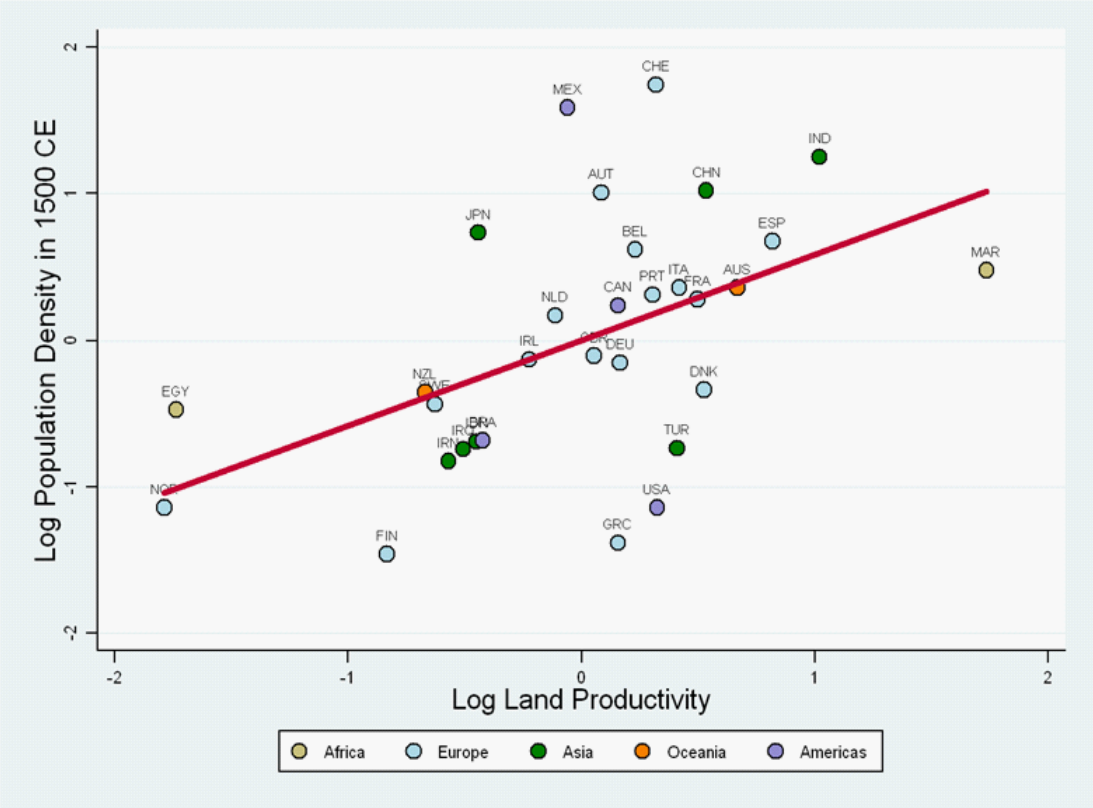
Conditional on land productivity, geographical factors, and continental fixed effects

Transition Timing and Income Per Capita in 1500 CE



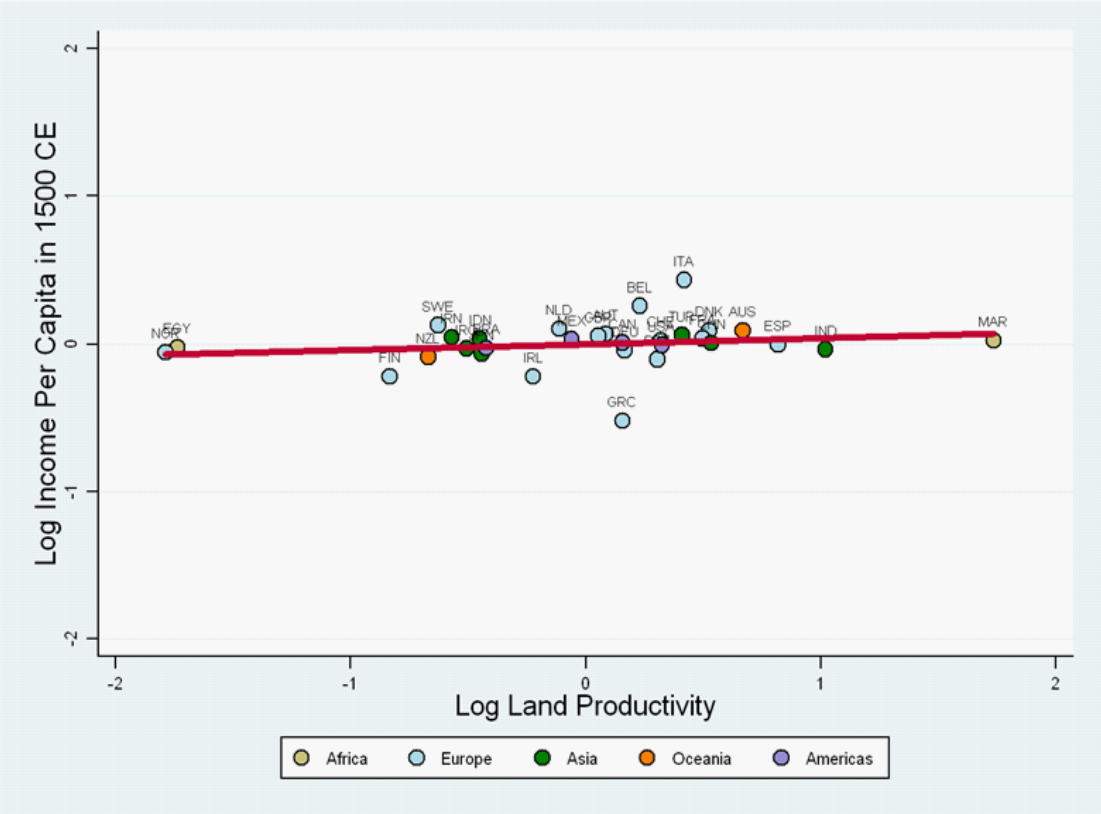
Conditional on land productivity, geographical factors, and continental fixed effects

Land Productivity and Population Density in 1500 CE



Conditional on transition timing, geographical factors, and continental fixed effects

Land Productivity and Income Per Capita in 1500 CE

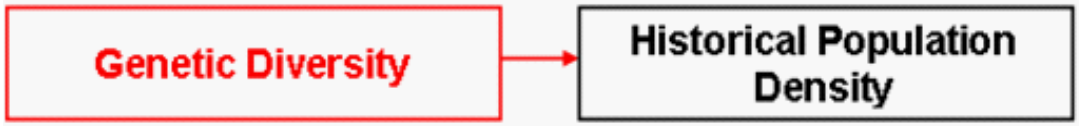


Conditional on transition timing, geographical factors, and continental fixed effects

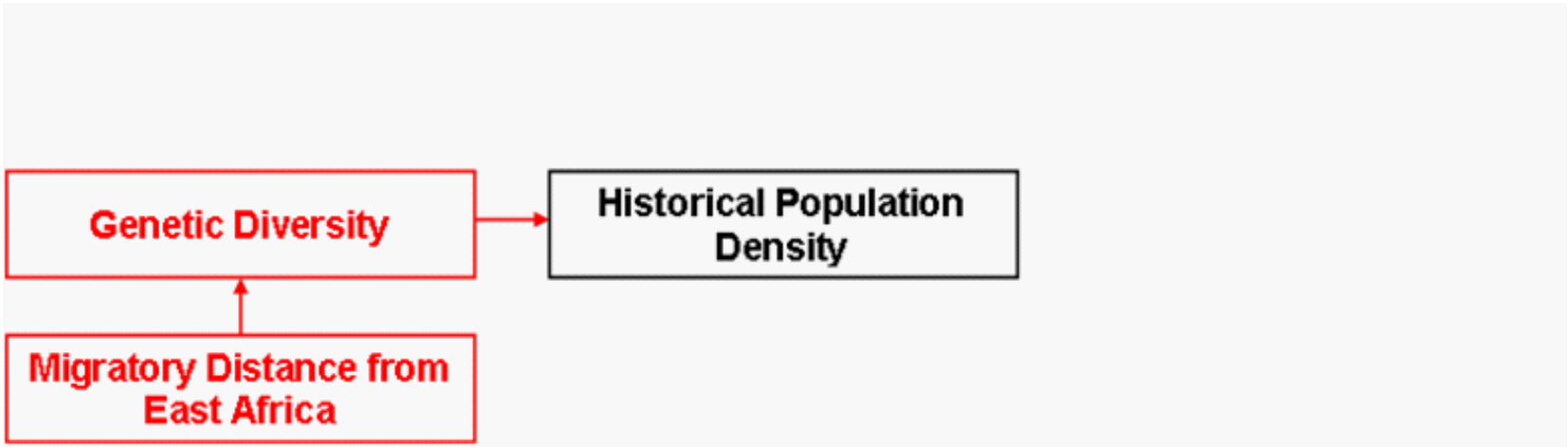
Summary of the Channels to be Examined

Historical Population Density

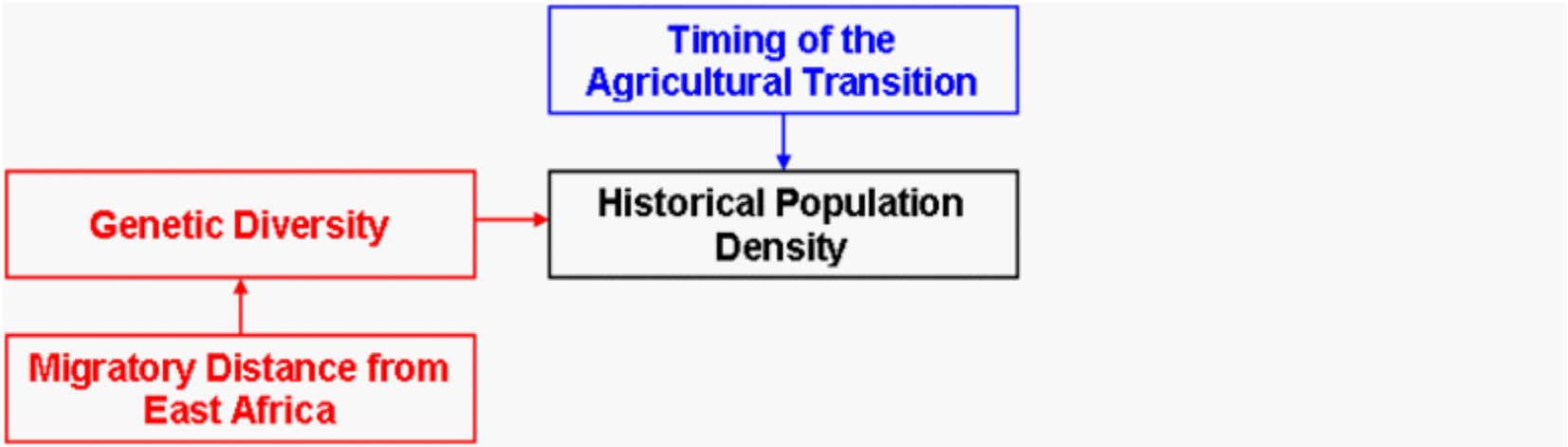
Summary of the Channels to be Examined



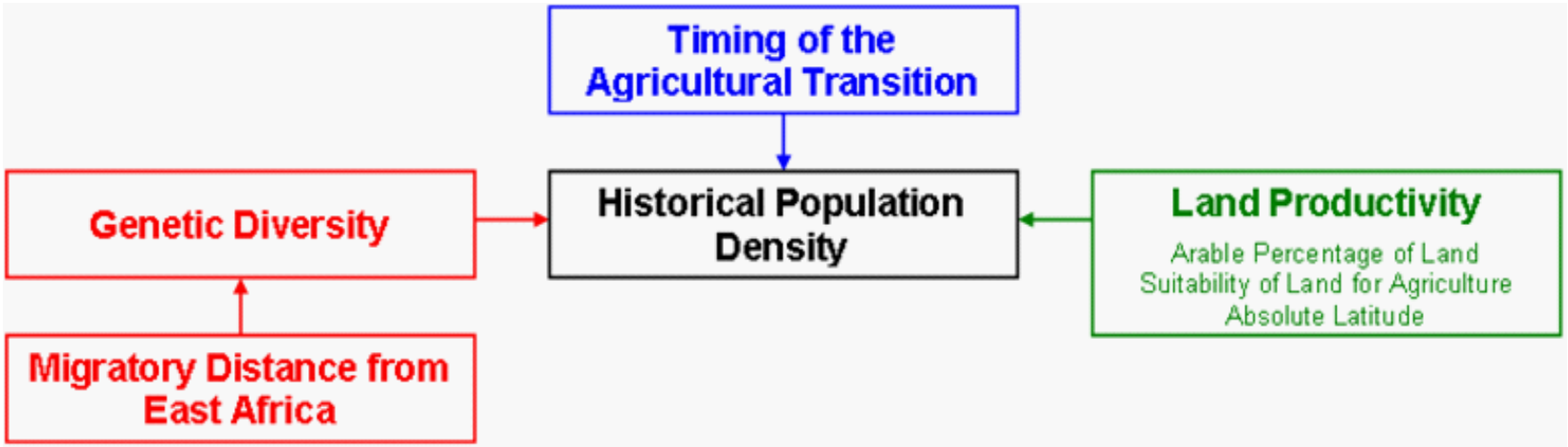
Summary of the Channels to be Examined



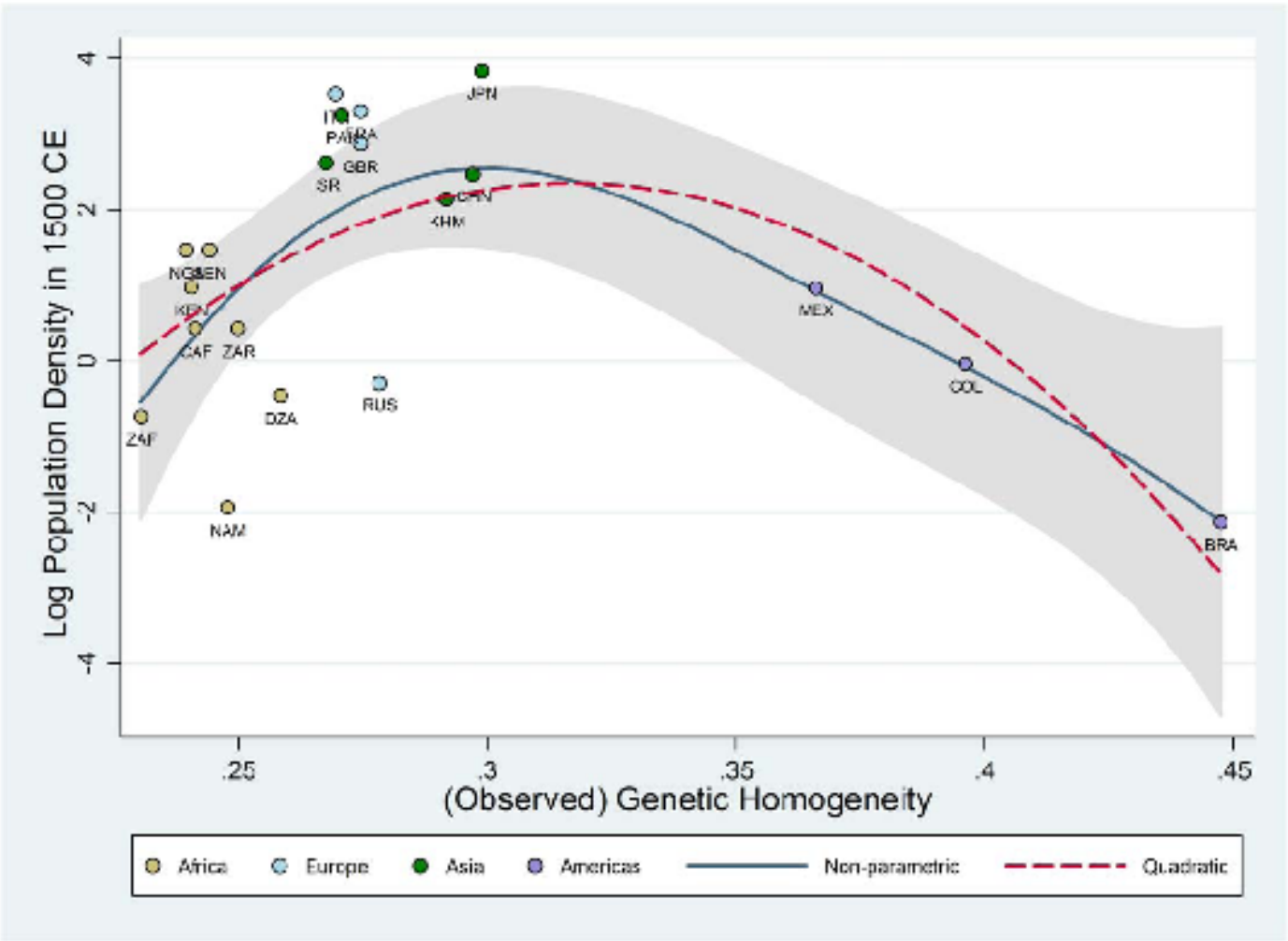
Summary of the Channels to be Examined



Summary of the Channels to be Examined



Observed Diversity and Comparative Development in 1500 CE: The Unconditional Relationship



Empirical Model I

- Use observed genetic diversity from the HGDP to test the hypothesis in a limited 21-country sample
- The empirical specification

$$\ln P_{it} = \beta_{0t} + \beta_{1t} G_i + \beta_{2t} G_i^2 + \beta_{3t} \ln T_{it} + \beta'_{4t} \ln X_i + \beta'_{5t} \ln \Delta_i + \varepsilon_{it}$$

where:

- $P_{it} \equiv$ population density of country i
- $G_i \equiv$ *actual* genetic diversity of country i
- $T_i \equiv$ timing of the Neolithic transition for country i
- $X_i \equiv$ vector of land productivity controls for country i
- $\Delta_i \equiv$ vector of continental dummies for country i
- $\varepsilon_{it} \equiv$ a country-year specific error term for country i

Empirical Model I

- Use observed genetic diversity from the HGDP to test the hypothesis in a limited 21-country sample
- The empirical specification

$$\ln P_{it} = \beta_{0t} + \beta_{1t} G_i + \beta_{2t} G_i^2 + \beta_{3t} \ln T_{it} + \beta'_{4t} \ln X_i + \beta'_{5t} \ln \Delta_i + \varepsilon_{it}$$

where:

- $P_{it} \equiv$ population density of country i
- $G_i \equiv$ *actual* genetic diversity of country i
- $T_i \equiv$ timing of the Neolithic transition for country i
- $X_i \equiv$ vector of land productivity controls for country i
- $\Delta_i \equiv$ vector of continental dummies for country i
- $\varepsilon_{it} \equiv$ a country-year specific error term for country i

Empirical Model I

- Use observed genetic diversity from the HGDP to test the hypothesis in a limited 21-country sample
- The empirical specification

$$\ln P_{it} = \beta_{0t} + \beta_{1t} G_i + \beta_{2t} G_i^2 + \beta_{3t} \ln T_{it} + \beta'_{4t} \ln X_i + \beta'_{5t} \ln \Delta_i + \varepsilon_{it}$$

where:

- $P_{it} \equiv$ population density of country i
- $G_i \equiv$ *actual* genetic diversity of country i
- $T_i \equiv$ timing of the Neolithic transition for country i
- $X_i \equiv$ vector of land productivity controls for country i
- $\Delta_i \equiv$ vector of continental dummies for country i
- $\varepsilon_{it} \equiv$ a country-year specific error term for country i

Empirical Model I

- Use observed genetic diversity from the HGDP to test the hypothesis in a limited 21-country sample
- The empirical specification

$$\ln P_{it} = \beta_{0t} + \beta_{1t} G_i + \beta_{2t} G_i^2 + \beta_{3t} \ln T_{it} + \beta'_{4t} \ln X_i + \beta'_{5t} \ln \Delta_i + \varepsilon_{it}$$

where:

- $P_{it} \equiv$ population density of country i
- $G_i \equiv$ *actual* genetic diversity of country i
- $T_i \equiv$ timing of the Neolithic transition for country i
- $X_i \equiv$ vector of land productivity controls for country i
- $\Delta_i \equiv$ vector of continental dummies for country i
- $\varepsilon_{it} \equiv$ a country-year specific error term for country i

Empirical Model I

- Use observed genetic diversity from the HGDP to test the hypothesis in a limited 21-country sample
- The empirical specification

$$\ln P_{it} = \beta_{0t} + \beta_{1t} G_i + \beta_{2t} G_i^2 + \beta_{3t} \ln T_{it} + \beta'_{4t} \ln X_i + \beta'_{5t} \ln \Delta_i + \varepsilon_{it}$$

where:

- $P_{it} \equiv$ population density of country i
- $G_i \equiv$ *actual* genetic diversity of country i
- $T_i \equiv$ timing of the Neolithic transition for country i
- $X_i \equiv$ vector of land productivity controls for country i
- $\Delta_i \equiv$ vector of continental dummies for country i
- $\varepsilon_{it} \equiv$ a country-year specific error term for country i

Empirical Model I

- Use observed genetic diversity from the HGDP to test the hypothesis in a limited 21-country sample
- The empirical specification

$$\ln P_{it} = \beta_{0t} + \beta_{1t} G_i + \beta_{2t} G_i^2 + \beta_{3t} \ln T_{it} + \beta'_{4t} \ln X_i + \beta'_{5t} \ln \Delta_i + \varepsilon_{it}$$

where:

- $P_{it} \equiv$ population density of country i
- $G_i \equiv$ *actual* genetic diversity of country i
- $T_i \equiv$ timing of the Neolithic transition for country i
- $X_i \equiv$ vector of land productivity controls for country i
- $\Delta_i \equiv$ vector of continental dummies for country i
- $\varepsilon_{it} \equiv$ a country-year specific error term for country i

Empirical Model I

- Use observed genetic diversity from the HGDP to test the hypothesis in a limited 21-country sample
- The empirical specification

$$\ln P_{it} = \beta_{0t} + \beta_{1t} G_i + \beta_{2t} G_i^2 + \beta_{3t} \ln T_{it} + \beta'_{4t} \ln X_i + \beta'_{5t} \ln \Delta_i + \varepsilon_{it}$$

where:

- $P_{it} \equiv$ population density of country i
- $G_i \equiv$ *actual* genetic diversity of country i
- $T_i \equiv$ timing of the Neolithic transition for country i
- $X_i \equiv$ vector of land productivity controls for country i
- $\Delta_i \equiv$ vector of continental dummies for country i
- $\varepsilon_{it} \equiv$ a country-year specific error term for country i

Empirical Model I

- Use observed genetic diversity from the HGDP to test the hypothesis in a limited 21-country sample
- The empirical specification

$$\ln P_{it} = \beta_{0t} + \beta_{1t} G_i + \beta_{2t} G_i^2 + \beta_{3t} \ln T_{it} + \beta'_{4t} \ln X_i + \beta'_{5t} \ln \Delta_i + \varepsilon_{it}$$

where:

- $P_{it} \equiv$ population density of country i
- $G_i \equiv$ *actual* genetic diversity of country i
- $T_i \equiv$ timing of the Neolithic transition for country i
- $X_i \equiv$ vector of land productivity controls for country i
- $\Delta_i \equiv$ vector of continental dummies for country i
- $\varepsilon_{it} \equiv$ a country-year specific error term for country i

Actual Diversity and Comparative Development in 1500 CE

	(1)	(2)	(3)	(4)	(5)
	Dependent Variable is Log Population Density in 1500 CE				
Genetic Diversity	413.51***			225.44***	203.82*
	(97.32)			(73.78)	(97.64)
Genetic Diversity Sqr.	-302.65***			-161.16**	-145.72*
	(73.34)			(56.16)	(80.41)
Log Transition Timing		2.40***		1.21***	1.14
		(0.27)		(0.37)	(0.66)
Log Arable % of Land			0.73**	0.52***	0.55*
			(0.28)	(0.17)	(0.26)
Log Absolute Latitude			0.15	-0.16	-0.13
			(0.18)	(0.13)	(0.17)
Log Agri. Suitability			0.73*	0.57*	0.59
			(0.38)	(0.29)	(0.33)
Optimal Diversity	0.683			0.699	0.699
	(0.008)			(0.015)	(0.055)
Continent Dummies	No	No	No	No	Yes
Observations	21	21	21	21	21
R-squared	0.42	0.54	0.57	0.89	0.90

Notes: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Migratory Distance, Genetic Diversity, and Comparative Development in 1500 CE

Genetic Diversity vs. Alternative Measures of Migratory Distance from East Africa

	(1)	(2)	(3)	(4)	(5)
	Dependent Variable is Log Population Density in 1500 CE				
Genetic Diversity	417.003***			300.978***	361.421**
	(90.909)			(76.371)	(121.429)
Genetic Diversity Sqr.	-306.218***			-241.755***	-268.515***
	(68.308)			(61.099)	(87.342)
Mig. Distance		0.463***		-0.003	
		(0.142)		(0.178)	
Mig. Distance Sqr.		-0.021***		-0.010	
		(0.006)		(0.009)	
Mobility Index			0.353**		0.051
			(0.127)		(0.154)
Mobility Index Sqr.			-0.012***		-0.003
			(0.004)		(0.006)
Observations	18	18	18	18	18
R-squared	0.43	0.30	0.30	0.47	0.43
P-value for:					
Joint Sig. of Diversity and its Sqr.				0.006	0.027
Joint Sig. of Distance and its Sqr.				0.320	
Joint Sig. of Mobility and its Sqr.					0.905

Notes: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Empirical Model II

- Use projected genetic diversity in an extended 145-country sample
- The empirical specification

$$\ln P_{it} = \beta_{0t} + \beta_{1t}\hat{G}_i + \beta_{2t}\hat{G}_i^2 + \beta_{3t}\ln T_{it} + \beta'_{4t}\ln X_i + \beta'_{5t}\ln \Delta_i + \varepsilon_{it}$$

where:

  - $P_{it} \equiv$ population density of country i in period t
  - $\hat{G}_i \equiv$ genetic diversity of country i *projected by migratory distance*
  - $T_i \equiv$ timing of the Neolithic transition for country i in period t
  - $X_i \equiv$ vector of land productivity controls for country i
  - $\Delta_i \equiv$ vector of continental dummies for country i
  - $\varepsilon_{it} \equiv$ a country-year specific error term for country i

Empirical Model II

- Use projected genetic diversity in an extended 145-country sample
- The empirical specification

$$\ln P_{it} = \beta_{0t} + \beta_{1t}\hat{G}_i + \beta_{2t}\hat{G}_i^2 + \beta_{3t}\ln T_{it} + \beta'_{4t}\ln X_i + \beta'_{5t}\ln \Delta_i + \varepsilon_{it}$$

where:

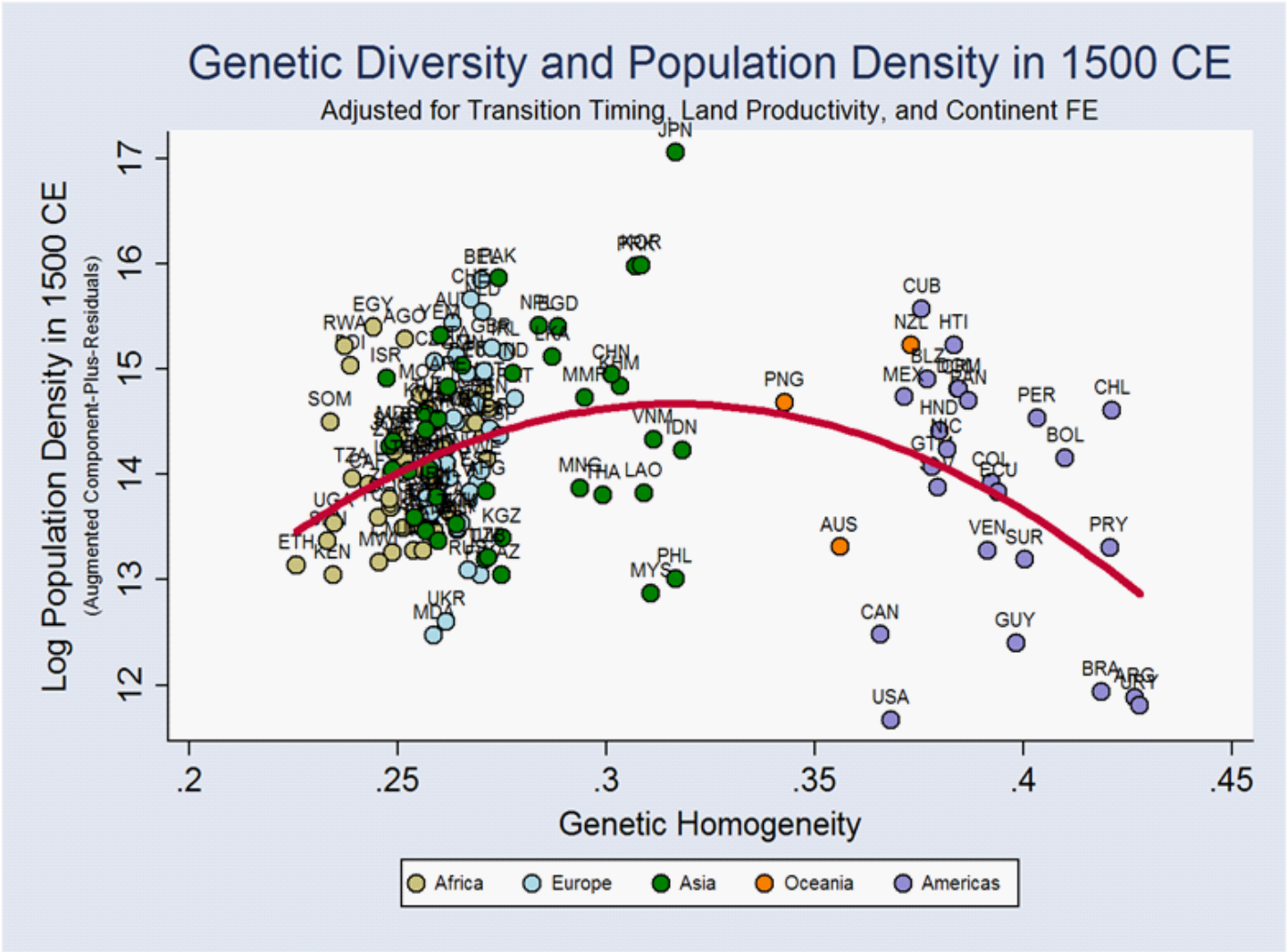
- $P_{it} \equiv$ population density of country i in period t
- $\hat{G}_i \equiv$ genetic diversity of country i *projected by migratory distance*
- $T_i \equiv$ timing of the Neolithic transition for country i in period t
- $X_i \equiv$ vector of land productivity controls for country i
- $\Delta_i \equiv$ vector of continental dummies for country i
- $\varepsilon_{it} \equiv$ a country-year specific error term for country i

Predicted Diversity and Comparative Development in 1500 CE

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable is Log Population Density in 1500 CE					
Pred. Diversity	250.99***		213.54***	203.02***	195.42***	199.73**
	(68.26)		(63.50)	(61.05)	(56.09)	(80.51)
Pred. Diversity Sqr.	-177.40***		-152.11***	-141.98***	-137.98***	-146.17***
	(50.22)		(46.65)	(44.83)	(40.84)	(56.26)
Log Transition Timing		1.29***	1.05***		1.16***	1.24***
		(0.18)	(0.19)		(0.15)	(0.24)
Log Arable % of Land				0.52***	0.40***	0.39***
				(0.12)	(0.09)	(0.10)
Log Absolute Latitude				-0.17*	-0.34***	-0.42***
				(0.09)	(0.09)	(0.12)
Log Agri. Suitability				0.19	0.31***	0.26***
				(0.12)	(0.10)	(0.10)
Optimal Diversity	0.707		0.702	0.715	0.708	0.683
	(0.021)		(0.025)	(0.110)	(0.051)	(0.110)
Continent Dummies	No	No	No	No	No	Yes
Observations	145	145	145	145	145	145
R-squared	0.22	0.26	0.38	0.50	0.67	0.69

Notes: Bootstrap standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Predicted Diversity and Comparative Development in 1500 CE



Interpretations - Diversity and Comparative Development in 1500

- Optimal GD $= 0.6832 \approx$ GD in Japan $= 0.6835$
- Increasing the diversity of the most homogeneous populations in South America by 11 percentage points to that of Japan, would have raised their population density in 1500 by a factor of 6
- Decreasing the diversity of to that of the most heterogeneous groups in East Africa, by 9 percentage points would have raised population density in 1500 by a factor of 3

Interpretations - Diversity and Comparative Development in 1500

- Optimal GD $= 0.6832 \approx$ GD in Japan $= 0.6835$
- Increasing the diversity of the most homogeneous populations in South America by 11 percentage points to that of Japan, would have raised their population density in 1500 by a factor of 6
- Decreasing the diversity of to that of the most heterogeneous groups in East Africa, by 9 percentage points would have raised population density in 1500 by a factor of 3

Interpretations - Diversity and Comparative Development in 1500

- Optimal GD $= 0.6832 \approx$ GD in Japan $= 0.6835$
- Increasing the diversity of the most homogeneous populations in South America by 11 percentage points to that of Japan, would have raised their population density in 1500 by a factor of 6
- Decreasing the diversity of to that of the most heterogeneous groups in East Africa, by 9 percentage points would have raised population density in 1500 by a factor of 3

Interpretation

- Increasing the diversity of the most homogeneous populations in South America by 1 percentage points, would have raised their population density in 1500 by 44%
- Decreasing the diversity of the most heterogeneous population in East Africa, by 1 percentage points would have raised population density in 1500 by 18%
- 1 percentage change in optimal level of diversity would have lowered population density in 1500 by 1.4%

Interpretation

- Increasing the diversity of the most homogeneous populations in South America by 1 percentage points, would have raised their population density in 1500 by 44%
- Decreasing the diversity of the most heterogeneous population in East Africa, by 1 percentage points would have raised population density in 1500 by 18%
- 1 percentage change in optimal level of diversity would have lowered population density in 1500 by 1.4%

Interpretation

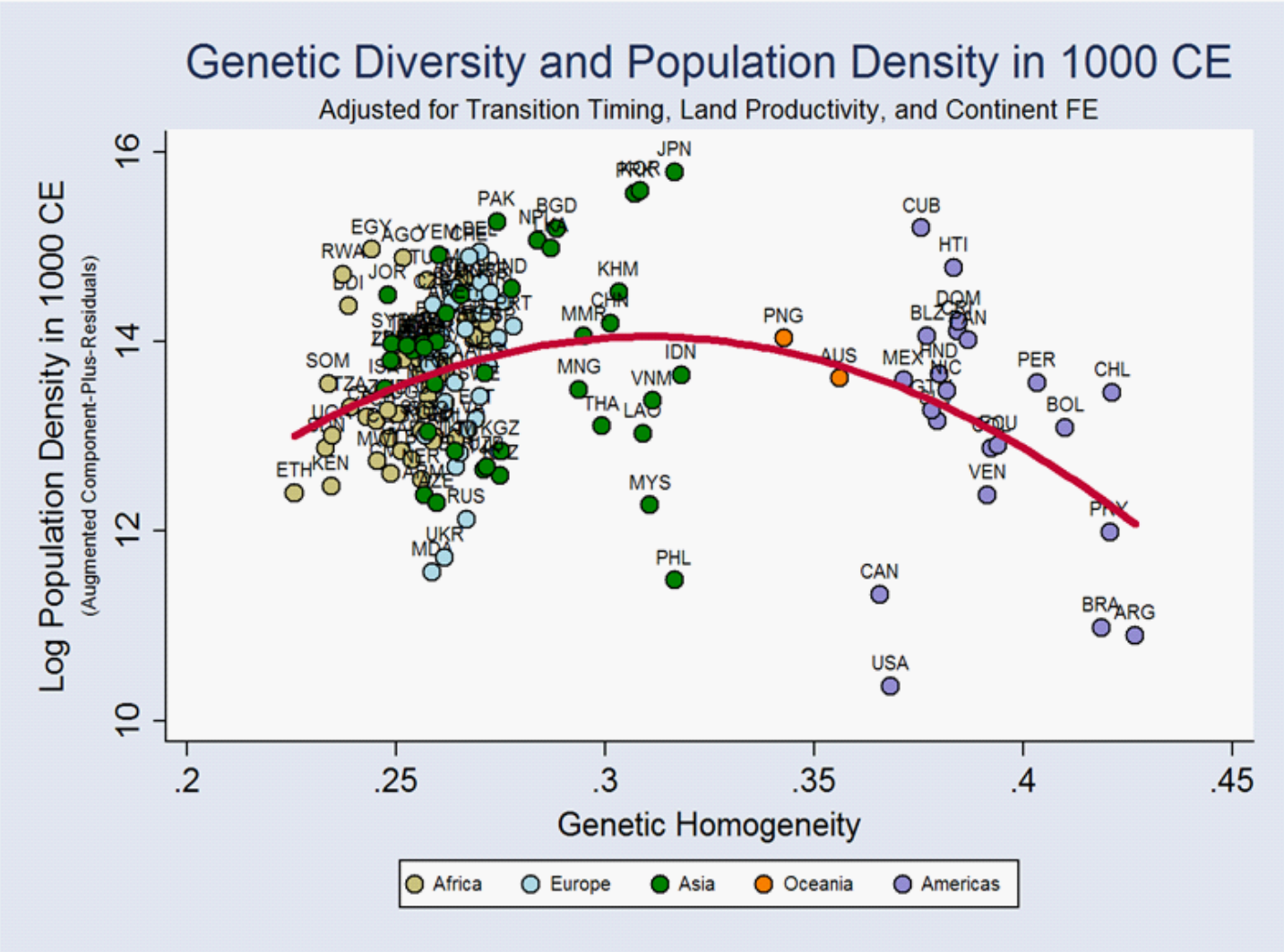
- Increasing the diversity of the most homogeneous populations in South America by 1 percentage points, would have raised their population density in 1500 by 44%
- Decreasing the diversity of the most heterogeneous population in East Africa, by 1 percentage points would have raised population density in 1500 by 18%
- 1 percentage change in optimal level of diversity would have lowered population density in 1500 by 1.4%

Predicted Diversity and Comparative Development in Earlier Periods

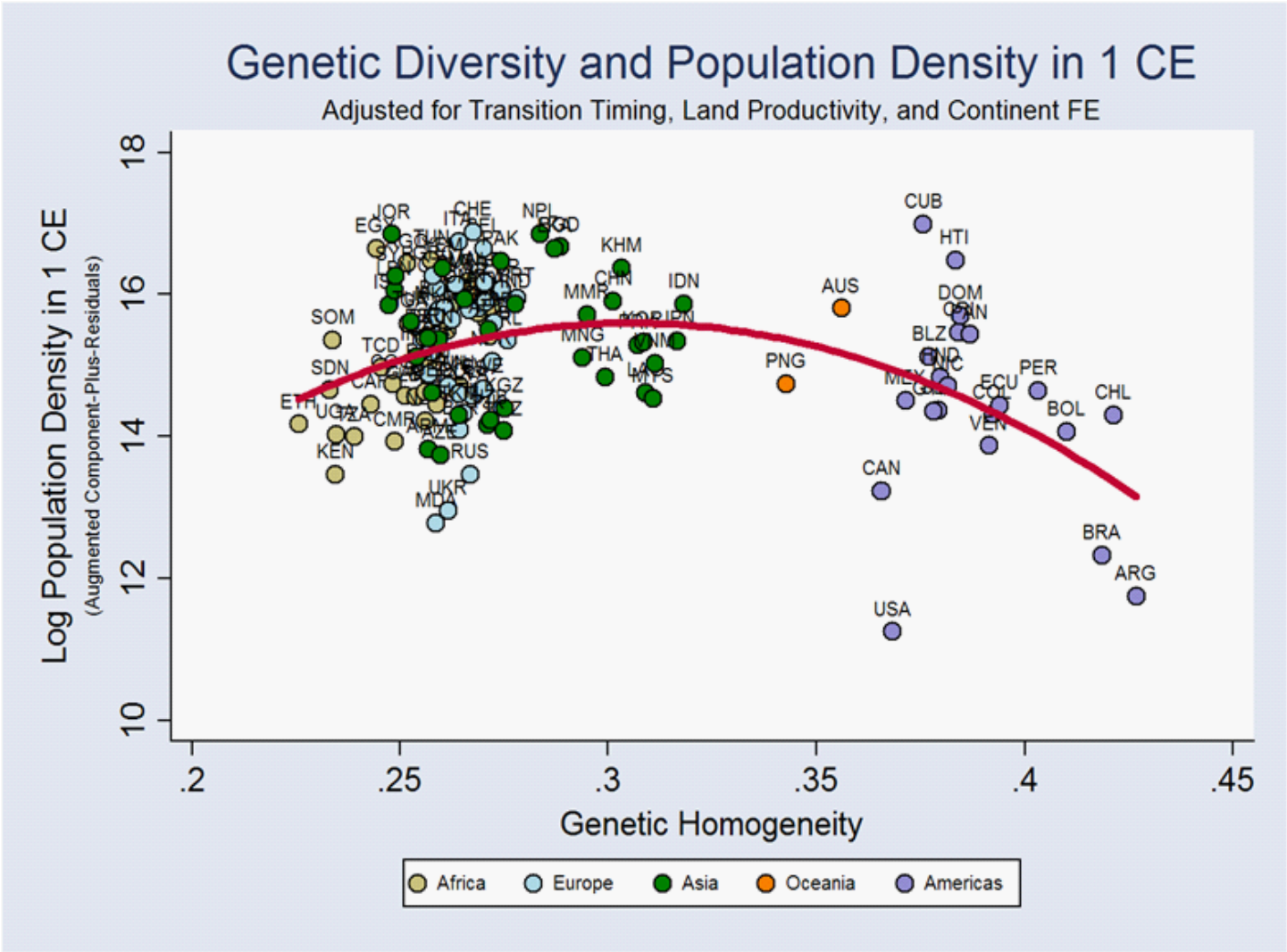
	(1)	(2)	(3)	(4)
	Dependent Variable is Log Population Density in			
	1000 CE	1000 CE	1 CE	1 CE
Pred. Diversity	154.91**	201.24**	134.77**	231.69**
	(62.39)	(95.58)	(63.45)	(115.83)
Pred. Diversity Sqr.	-109.81**	-145.89**	-96.25**	-166.86**
	(45.70)	(66.79)	(46.49)	(81.13)
Log Transition Timing	1.37***	1.60***	1.66***	2.13***
	(0.15)	(0.27)	(0.21)	(0.44)
Log Arable % of Land	0.37***	0.37***	0.31***	0.35***
	(0.10)	(0.12)	(0.12)	(0.13)
Log Absolute Latitude	-0.38***	-0.37***	-0.12	-0.12
	(0.10)	(0.14)	(0.12)	(0.13)
Log Agri. Suitability	0.21**	0.19*	0.24*	0.21*
	(0.10)	(0.11)	(0.12)	(0.12)
Optimal Diversity	0.705	0.690	0.705	0.694
	(0.108)	(0.293)	(0.188)	(0.194)
Continent Dummies	No	Yes	No	Yes
Observations	140	140	126	126
R-squared	0.61	0.62	0.59	0.61

Notes: Bootstrap standard errors in parentheses;
*** p<0.01, ** p<0.05, * p<0.1

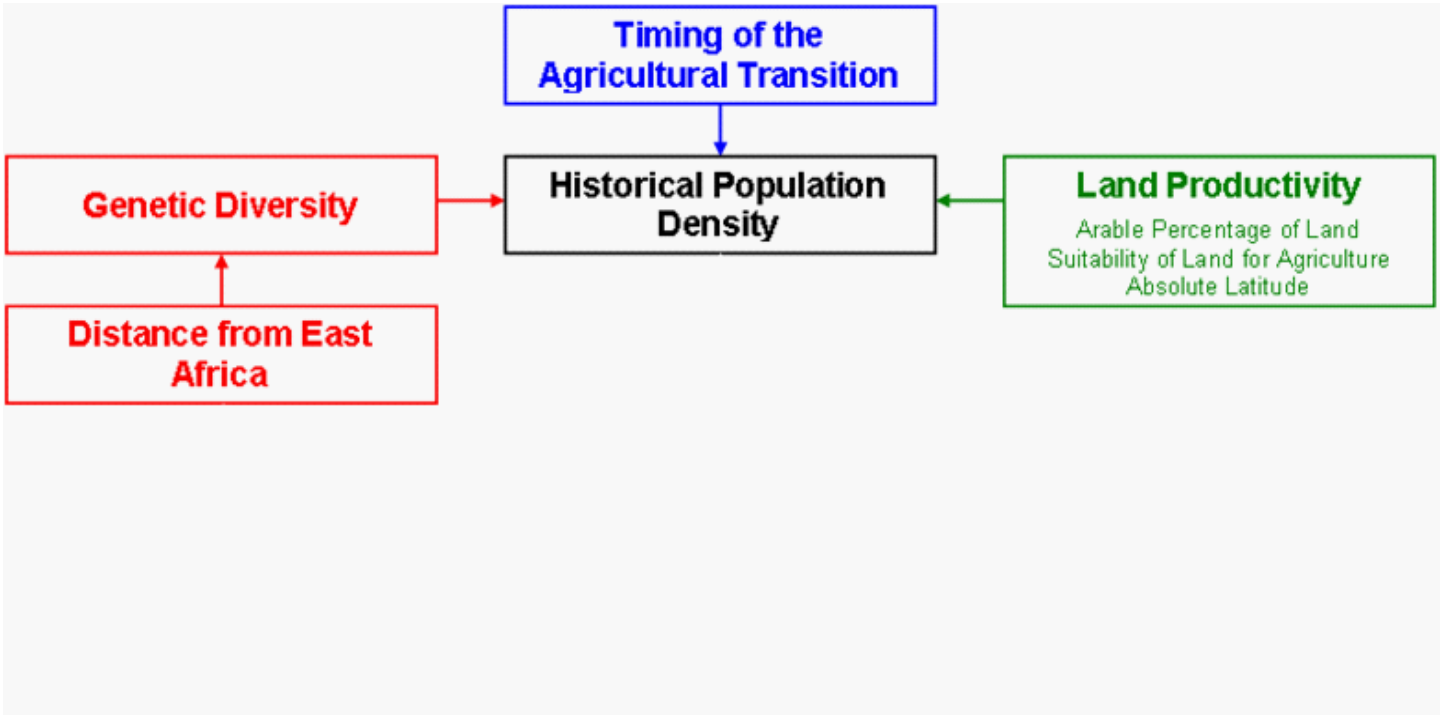
Predicted Diversity and Comparative Development in 1000 CE



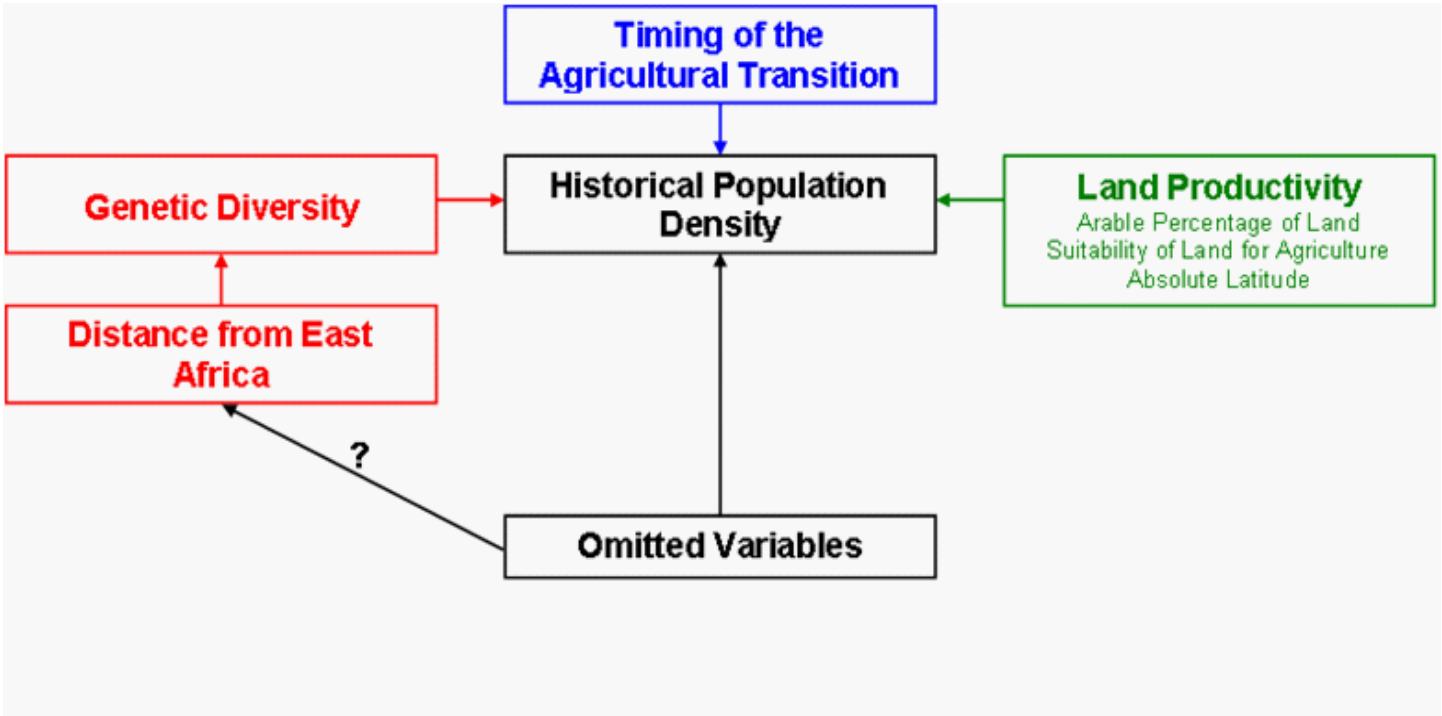
Predicted Diversity and Comparative Development in 1 CE



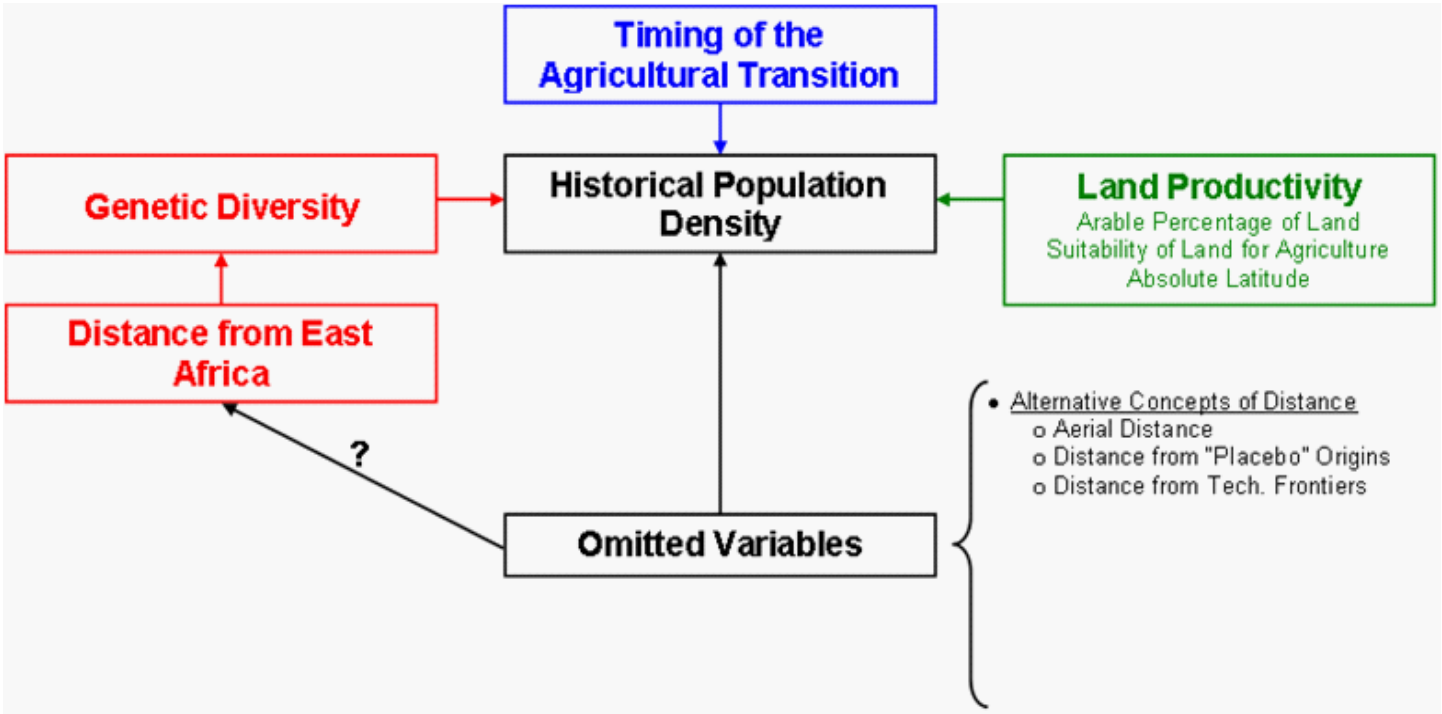
Robustness Analysis



Robustness Analysis



Robustness Analysis



Robustness to Aerial Distance and Migratory Distances from "Placebo" Origins

	(1)	(2)	(3)	(4)	(5)
	Dependent Variable is Log Population Density in 1500 CE				
Distance calculated from:	Addis Ababa	Addis Ababa	London	Tokyo	Mexico City
Migratory Distance	0.138**		-0.040	0.052	-0.063
	(0.061)		(0.063)	(0.145)	(0.099)
Migratory Distance Sqr.	-0.008***		-0.002	-0.006	0.005
	(0.002)		(0.002)	(0.007)	(0.004)
Aerial Distance		-0.008			
		(0.106)			
Aerial Distance Sqr.		-0.005			
		(0.006)			
Log Transition Timing	1.160***	1.158***	1.003***	1.047***	1.619***
	(0.144)	(0.138)	(0.164)	(0.225)	(0.277)
Log Arable % of Land	0.401***	0.488***	0.357***	0.532***	0.493***
	(0.091)	(0.102)	(0.092)	(0.089)	(0.094)
Log Absolute Latitude	-0.342***	-0.263***	-0.358***	-0.334***	-0.239***
	(0.091)	(0.097)	(0.112)	(0.099)	(0.083)
Log Agri. Suitability	0.305***	0.254**	0.344***	0.178**	0.261***
	(0.091)	(0.102)	(0.092)	(0.080)	(0.092)
Observations	145	145	145	145	145
R-squared	0.67	0.59	0.67	0.59	0.63

Notes: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Robustness to Distance from Technological Frontiers

- Identify historical technological frontiers using urban development as a proxy for technological advancement (Kuznets (1968), Bairoch (1988), Chandler (1987))
- For each continent and each historical period, select the two largest cities (belonging to different sociopolitical entities) to construct a set of regional technological frontiers
- For each period examined, choose the closest from among this set as the relevant frontier for the observation in question

Robustness to Distance from Technological Frontiers

- Identify historical technological frontiers using urban development as a proxy for technological advancement (Kuznets (1968), Bairoch (1988), Chandler (1987))
- For each continent and each historical period, select the two largest cities (belonging to different sociopolitical entities) to construct a set of regional technological frontiers
- For each period examined, choose the closest from among this set as the relevant frontier for the observation in question

Robustness to Distance from Technological Frontiers

- Identify historical technological frontiers using urban development as a proxy for technological advancement (Kuznets (1968), Bairoch (1988), Chandler (1987))
- For each continent and each historical period, select the two largest cities (belonging to different sociopolitical entities) to construct a set of regional technological frontiers
- For each period examined, choose the closest from among this set as the relevant frontier for the observation in question

Regional Technological Frontiers

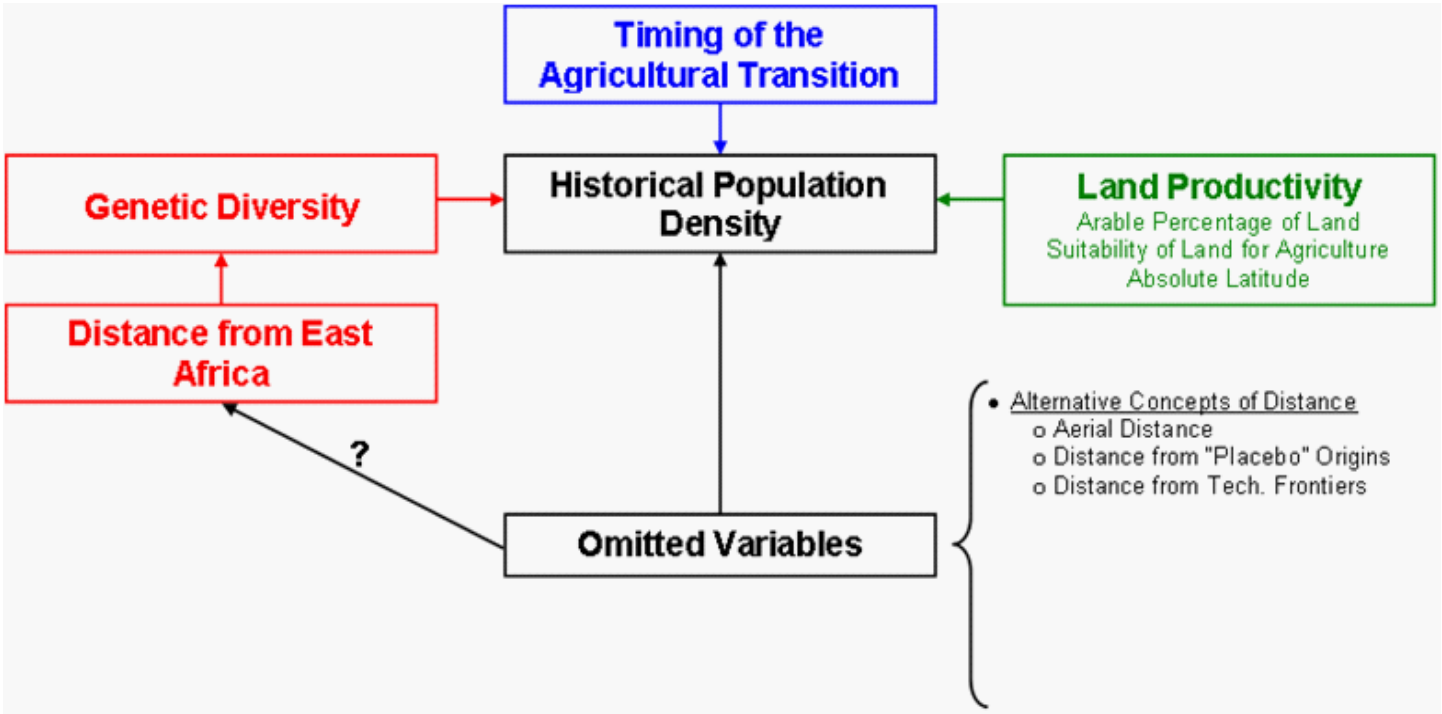
City and Modern Location	Continent	Sociopolitical Entity	Relevant Period
Cairo, Egypt	Africa	Mamluk Sultanate	1500 CE
Fez, Morocco	Africa	Marinid Kingdom of Fez	1500 CE
London, UK	Europe	Tudor Dynasty	1500 CE
Paris, France	Europe	Valois-Orléans Dynasty	1500 CE
Constantinople, Turkey	Asia	Ottoman Empire	1500 CE
Peking, China	Asia	Ming Dynasty	1500 CE
Tenochtitlan, Mexico	Americas	Aztec Civilization	1500 CE
Cuzco, Peru	Americas	Inca Civilization	1500 CE
Cairo, Egypt	Africa	Fatimid Caliphate	1000 CE
Kairwan, Tunisia	Africa	Berber Zirite Dynasty	1000 CE
Constantinople, Turkey	Europe	Byzantine Empire	1000 CE
Cordoba, Spain	Europe	Caliphate of Cordoba	1000 CE
Baghdad, Iraq	Asia	Abbasid Caliphate	1000 CE
Kaifeng, China	Asia	Song Dynasty	1000 CE
Tollan, Mexico	Americas	Classic Maya Civilization	1000 CE
Huari, Peru	Americas	Huari Culture	1000 CE
Alexandria, Egypt	Africa	Roman Empire	1 CE
Carthage, Tunisia	Africa	Roman Empire	1 CE
Athens, Greece	Europe	Roman Empire	1 CE
Rome, Italy	Europe	Roman Empire	1 CE
Luoyang, China	Asia	Han Dynasty	1 CE
Seleucia, Iraq	Asia	Seleucid Dynasty	1 CE
Teotihuacán, Mexico	Americas	Pre-classic Maya Civilization	1 CE
Cahuachi, Peru	Americas	Nazca Culture	1 CE

Robustness to Distance from Technological Frontiers

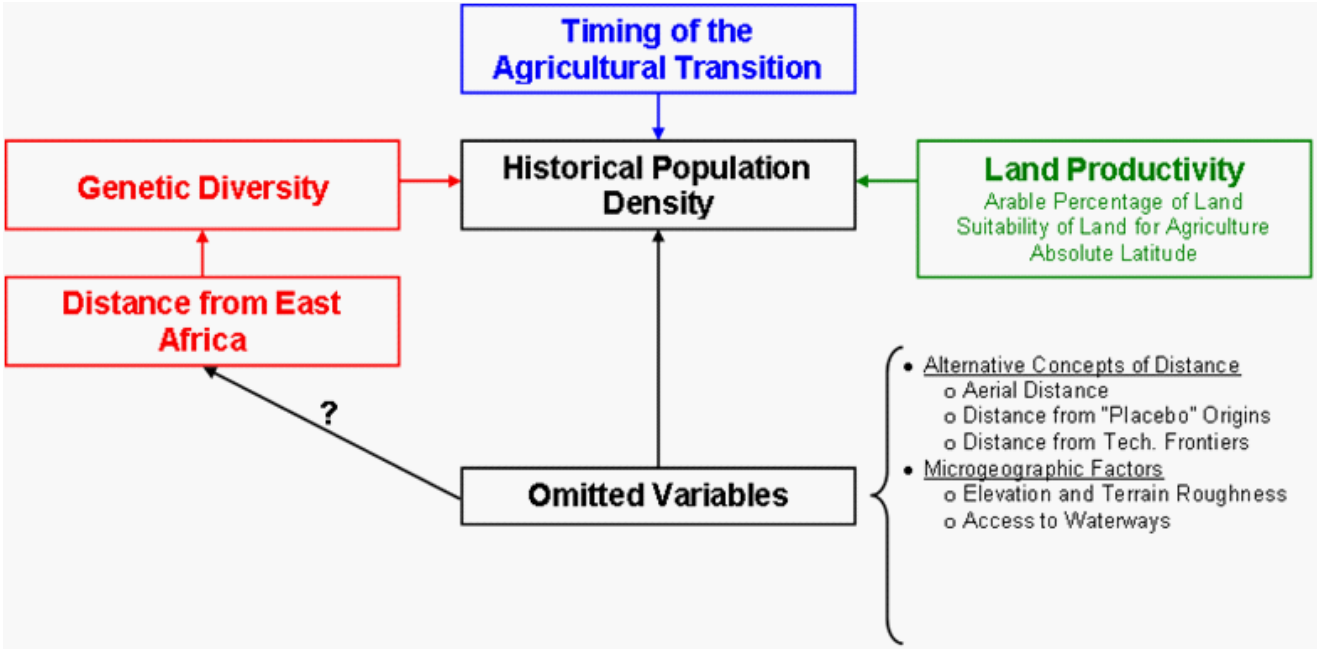
	(1)	(2)	(3)	(4)	(5)	(6)
	Log Population Density in 1500 CE		Log Population Density in 1000 CE		Log Population Density in 1 CE	
Pred. Diversity	156.74**		183.77**		215.86**	
	(77.98)		(91.20)		(106.50)	
Pred. Diversity Sqr.	-114.63**		-134.61**		-157.72**	
	(54.67)		(63.65)		(74.82)	
Log Transition Timing	Yes		Yes		Yes	
Land Prod. Controls	Yes		Yes		Yes	
Log Dist. to Regional Frontier in 1500 CE	-0.19***					
	(0.07)					
Log Dist. to Regional Frontier in 1000 CE			-0.23**			
			(0.11)			
Log Dist. to Regional Frontier in 1 CE					-0.30***	
					(0.10)	
Optimal Diversity	0.684		0.683		0.684	
	(0.169)		(0.218)		(0.266)	
Continent Dummies	Yes		Yes		Yes	
Observations	145		140		126	
R-squared	0.72		0.64		0.66	

Notes: Bootstrap standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Robustness Analysis



Robustness Analysis

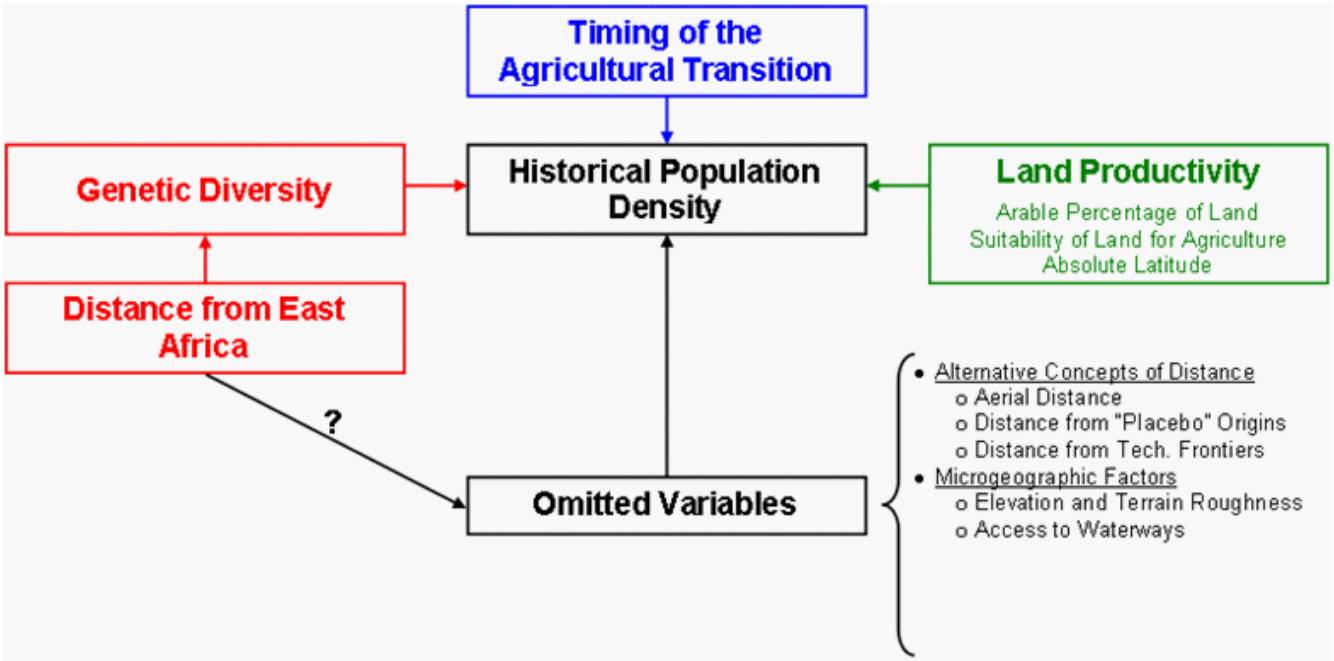


Robustness to Microgeography

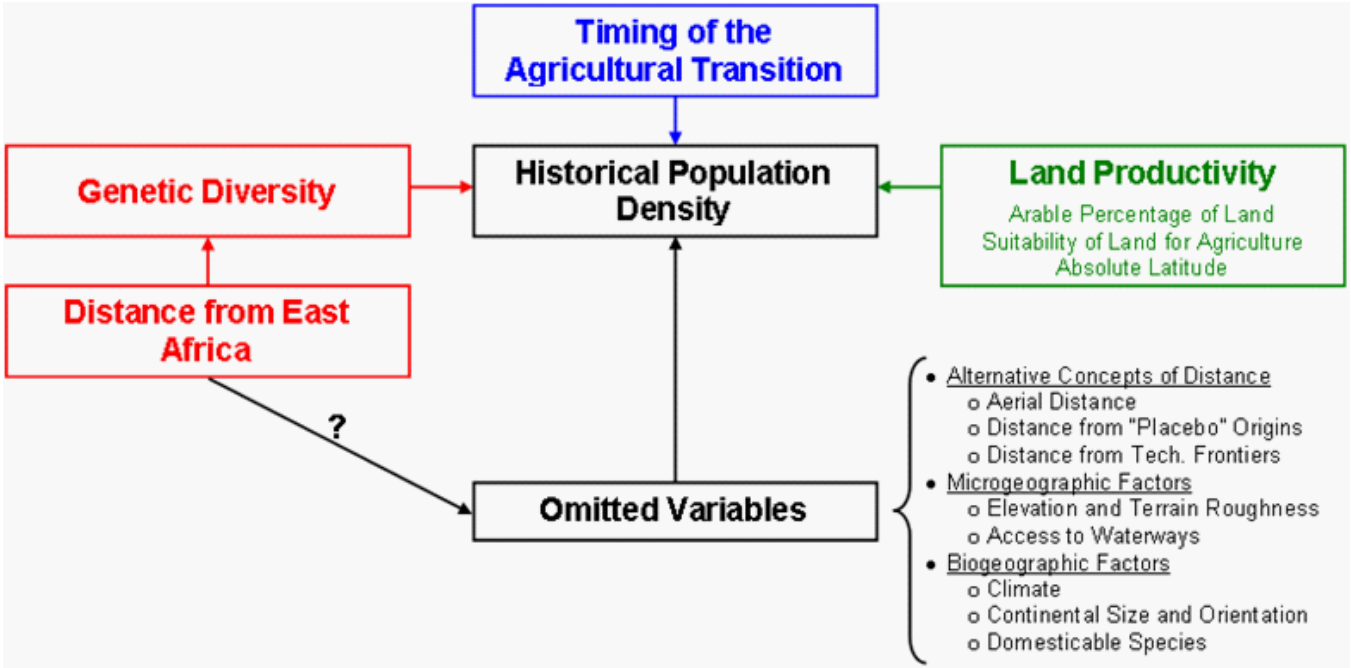
	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable is Log Population Density in 1500 CE					
Pred. Diversity	159.92***		153.20***	157.07**	150.02***	157.06**
	(56.00)		(53.39)	(78.82)	(49.36)	(68.61)
Pred. Diversity Sqr.	-110.39***		-105.33***	-112.78**	-102.76***	-114.99**
	(41.08)		(39.11)	(55.48)	(36.23)	(48.26)
Log Transition Timing	Yes		Yes	Yes	Yes	Yes
Land Prod. Controls	Yes		Yes	Yes	Yes	Yes
Mean Elevation	-0.48**				0.51*	0.50*
	(0.23)				(0.27)	(0.27)
Roughness	5.15***				3.09*	4.08**
	(1.77)				(1.74)	(1.84)
Roughness Sqr.	-7.05**				-7.05**	-7.63***
	(3.11)				(2.96)	(2.91)
Distance to Nearest Waterway			-0.49***	-0.44**	-0.47**	-0.39**
			(0.18)	(0.18)	(0.18)	(0.18)
% Land within 100 km of Waterway			0.70**	0.73**	1.11***	1.18***
			(0.28)	(0.31)	(0.29)	(0.29)
Optimal Diversity	0.724***		0.727***	0.696***	0.730***	0.683
	(0.201)		(0.190)	(0.187)	(0.229)	(0.095)
Continent Dummies	No		No	Yes	No	Yes
Observations	145		145	145	145	145
R-squared	0.69		0.74	0.75	0.76	0.78

Notes: Bootstrap standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Robustness Analysis



Robustness Analysis



Robustness to Biogeography

Robustness to Ultimate Determinants in the Diamond Hypothesis

	(1)	(2)	(3)	(4)	(5)
	Dependent Variable is Log Population Density in 1500 CE				
Pred. Diversity	216.85***	252.08***	174.41***	212.12***	274.92***
	(62.06)	(70.81)	(62.75)	(72.13)	(72.12)
Pred. Diversity Sqr.	-154.75***	-180.65***	-125.14***	-151.58***	-197.12***
	(45.19)	(51.89)	(45.72)	(52.79)	(52.40)
Log Transition Timing	1.30***				1.16***
	(0.16)				(0.31)
Land Prod. Controls	Yes	Yes	Yes	Yes	Yes
Climate		0.62***		0.42	0.37*
		(0.14)		(0.27)	(0.22)
Orientation of Axis		0.28		0.04	-0.17
		(0.33)		(0.30)	(0.27)
Size of Continent		-0.01		-0.01	-0.01
		(0.02)		(0.01)	(0.01)
Domesticable Plants			0.02	-0.01	0.00
			(0.02)	(0.02)	(0.02)
Domesticable Animals			0.15**	0.12	-0.01
			(0.06)	(0.07)	(0.07)
Optimal Diversity	0.701	0.698	0.697	0.700	0.697
	(0.123)	(0.016)	(0.159)	(0.045)	(0.041)
Observations	96	96	96	96	96
R-squared	0.74	0.70	0.70	0.72	0.78

Notes: Bootstrap standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Genetic Diversity Index in 2000

Account for:

- Genetic *diversity* among the ancestral populations of each country
- Genetic *distance* between the ancestral populations of each country
- Ethnic composition of each country (post-Colonial population flows)

Genetic Diversity Index in 2000

Account for:

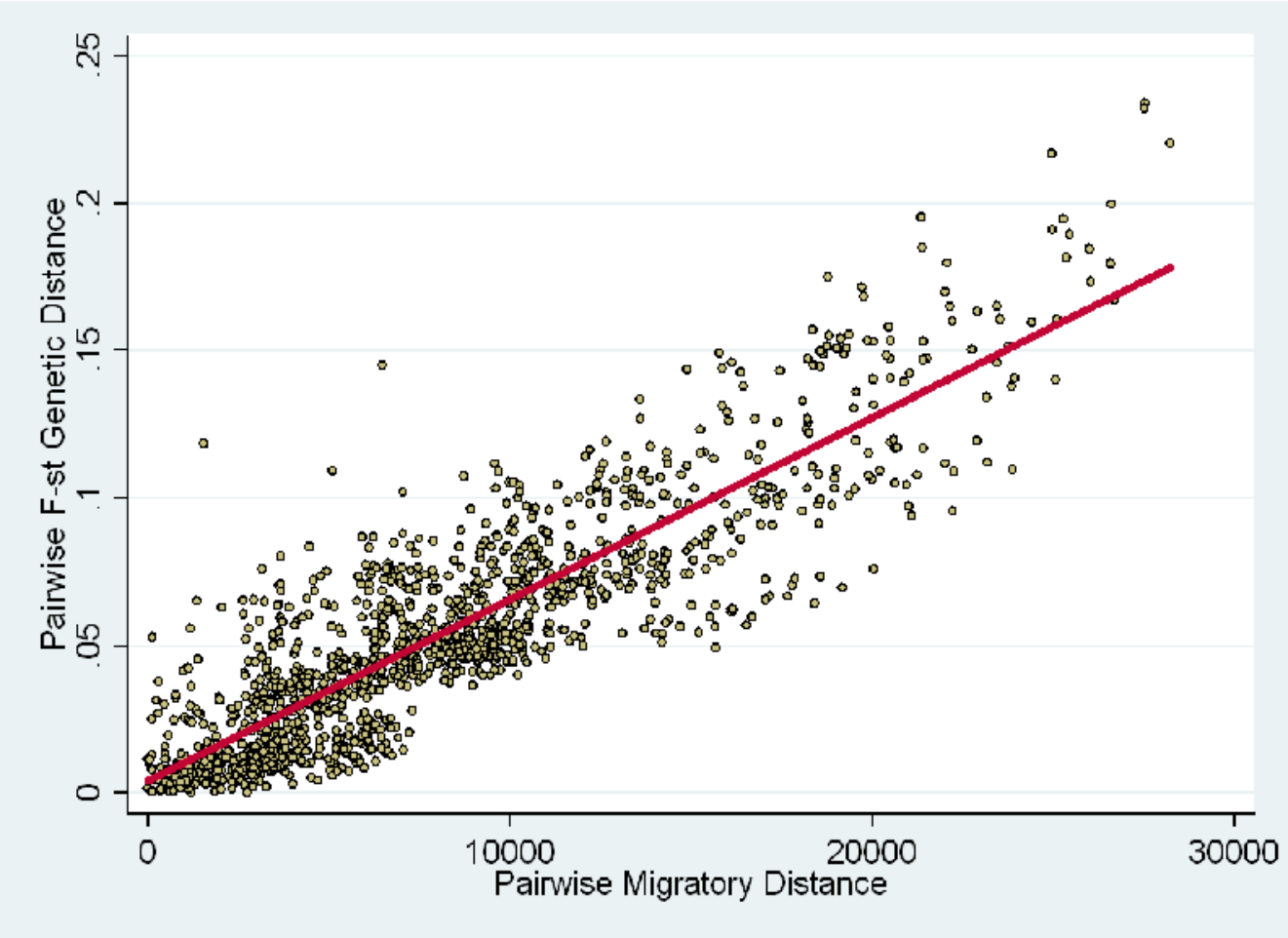
- Genetic *diversity* among the ancestral populations of each country
- Genetic *distance* between the ancestral populations of each country
- Ethnic composition of each country (post-Colonial population flows)

Genetic Diversity Index in 2000

Account for:

- Genetic *diversity* among the ancestral populations of each country
- Genetic *distance* between the ancestral populations of each country
- Ethnic composition of each country (post-Colonial population flows)

Projection of Genetic Distances



Empirical Model II

- The empirical specification

$$\ln y_i = \gamma_0 + \gamma_1 \hat{G}_i + \gamma_2 \hat{G}_i^2 + \gamma_3 \ln T_i + \gamma_4' \ln X_i + \gamma_5' \ln \Lambda_i + \gamma_6 \ln \Gamma_i + \eta_i$$

where:

- $y_i \equiv$ Income per capita of country i in the year 2000
- $\hat{G}_i \equiv$ Index of contemporary population diversity in country i
- $T_i \equiv$ timing of the Neolithic transition for country i
- $X_i \equiv$ vector of land productivity controls for country i
- $\Lambda_i \equiv$ vector of institutional and cultural controls for country i
- $\Gamma_i \equiv$ vector of additional geographical controls for country i

Empirical Model II

- The empirical specification

$$\ln y_i = \gamma_0 + \gamma_1 \hat{G}_i + \gamma_2 \hat{G}_i^2 + \gamma_3 \ln T_i + \gamma_4' \ln X_i + \gamma_5' \ln \Lambda_i + \gamma_6 \ln \Gamma_i + \eta_i$$

where:

  - $y_i \equiv$ Income per capita of country i in the year 2000
  - $\hat{G}_i \equiv$ Index of contemporary population diversity in country i
  - $T_i \equiv$ timing of the Neolithic transition for country i
  - $X_i \equiv$ vector of land productivity controls for country i
  - $\Lambda_i \equiv$ vector of institutional and cultural controls for country i
  - $\Gamma_i \equiv$ vector of additional geographical controls for country i

Empirical Model II

- The empirical specification

$$\ln y_i = \gamma_0 + \gamma_1 \hat{G}_i + \gamma_2 \hat{G}_i^2 + \gamma_3 \ln T_i + \gamma_4' \ln X_i + \gamma_5' \ln \Lambda_i + \gamma_6 \ln \Gamma_i + \eta_i$$

where:

  - $y_i \equiv$ Income per capita of country i in the year 2000
  - $\hat{G}_i \equiv$ Index of contemporary population diversity in country i
  - $T_i \equiv$ timing of the Neolithic transition for country i
  - $X_i \equiv$ vector of land productivity controls for country i
  - $\Lambda_i \equiv$ vector of institutional and cultural controls for country i
  - $\Gamma_i \equiv$ vector of additional geographical controls for country i

Empirical Model II

- The empirical specification

$$\ln y_i = \gamma_0 + \gamma_1 \hat{G}_i + \gamma_2 \hat{G}_i^2 + \gamma_3 \ln T_i + \gamma_4' \ln X_i + \gamma_5' \ln \Lambda_i + \gamma_6 \ln \Gamma_i + \eta_i$$

where:

  - $y_i \equiv$ Income per capita of country i in the year 2000
  - $\hat{G}_i \equiv$ Index of contemporary population diversity in country i
  - $T_i \equiv$ timing of the Neolithic transition for country i
  - $X_i \equiv$ vector of land productivity controls for country i
  - $\Lambda_i \equiv$ vector of institutional and cultural controls for country i
  - $\Gamma_i \equiv$ vector of additional geographical controls for country i

Empirical Model II

- The empirical specification

$$\ln y_i = \gamma_0 + \gamma_1 \hat{G}_i + \gamma_2 \hat{G}_i^2 + \gamma_3 \ln T_i + \gamma_4' \ln X_i + \gamma_5' \ln \Lambda_i + \gamma_6 \ln \Gamma_i + \eta_i$$

where:

  - $y_i \equiv$ Income per capita of country i in the year 2000
  - $\hat{G}_i \equiv$ Index of contemporary population diversity in country i
  - $T_i \equiv$ timing of the Neolithic transition for country i
  - $X_i \equiv$ vector of land productivity controls for country i
  - $\Lambda_i \equiv$ vector of institutional and cultural controls for country i
  - $\Gamma_i \equiv$ vector of additional geographical controls for country i

Empirical Model II

- The empirical specification

$$\ln y_i = \gamma_0 + \gamma_1 \hat{G}_i + \gamma_2 \hat{G}_i^2 + \gamma_3 \ln T_i + \gamma_4' \ln X_i + \gamma_5' \ln \Lambda_i + \gamma_6 \ln \Gamma_i + \eta_i$$

where:

  - $y_i \equiv$ Income per capita of country i in the year 2000
  - $\hat{G}_i \equiv$ Index of contemporary population diversity in country i
  - $T_i \equiv$ timing of the Neolithic transition for country i
  - $X_i \equiv$ vector of land productivity controls for country i
  - $\Lambda_i \equiv$ vector of institutional and cultural controls for country i
  - $\Gamma_i \equiv$ vector of additional geographical controls for country i

Empirical Model II

- The empirical specification

$$\ln y_i = \gamma_0 + \gamma_1 \hat{G}_i + \gamma_2 \hat{G}_i^2 + \gamma_3 \ln T_i + \gamma_4' \ln X_i + \gamma_5' \ln \Lambda_i + \gamma_6 \ln \Gamma_i + \eta_i$$

where:

  - $y_i \equiv$ Income per capita of country i in the year 2000
  - $\hat{G}_i \equiv$ Index of contemporary population diversity in country i
  - $T_i \equiv$ timing of the Neolithic transition for country i
  - $X_i \equiv$ vector of land productivity controls for country i
  - $\Lambda_i \equiv$ vector of institutional and cultural controls for country i
  - $\Gamma_i \equiv$ vector of additional geographical controls for country i

Diversity and Economic Development in 2000 and 1500 CE

	(1)	(2)	(3)	(4)
	Dependent Variable is:			
	Log Income Per Capita in 2000 CE			Log Population Density in 1500 CE
Adj. Diversity	204.610**	237.238***	244.960***	
	(88.466)	(86.278)	(85.454)	
Adj. Diversity Sqr.	-143.437**	-166.507***	-171.364***	
	(62.545)	(61.363)	(60.843)	
Unadj. Diversity				198.587**
				(79.110)
Unadj. Diversity Sqr.				-145.320***
				(55.472)
Log Adj. Trans. Timing		0.061	0.002	
		(0.262)	(0.305)	
Log Transition Timing	-0.151			1.238***
	(0.186)			(0.230)
Log Arable % of Land	-0.110	-0.119	-0.137	0.378***
	(0.100)	(0.107)	(0.111)	(0.100)
Log Absolute Latitude	0.164	0.172	0.192	-0.423***
	(0.125)	(0.119)	(0.143)	(0.124)
Log Agri. Suitability	-0.193**	-0.177*	-0.189*	0.264***
	(0.095)	(0.102)	(0.102)	(0.096)
Log Population Density in 1500 CE			0.047	
			(0.097)	
Optimal Diversity	0.713	0.712	0.715	0.683
	(0.100)	(0.036)	(0.118)	(0.095)
Continent Dummies	Yes	Yes	Yes	Yes
Observations	143	143	143	143
R-squared	0.57	0.57	0.57	0.68

Notes: Bootstrap standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Diversity and Comparative Development in 2000 CE

Genetic Diversity, Institutions, and Ethnic Fractionalization

	(1)	(2)	(3)	(4)
	Dependent Variable is Log Income Per Capita in 2000 CE			
Adj. Genetic Diversity	315.282***	225.858***	219.453***	219.813***
	(84.215)	(67.669)	(67.558)	(68.965)
Adj. Genetic Diversity Sqr.	-220.980***	-155.826***	-151.489***	-151.734***
	(59.562)	(47.962)	(47.792)	(48.765)
Log Adj. Transition Timing	-0.273	-0.092	0.032	0.032
	(0.269)	(0.200)	(0.189)	(0.191)
Log Arable % of Land	-0.218***	-0.159***	-0.171***	-0.172***
	(0.061)	(0.049)	(0.049)	(0.049)
Log Absolute Latitude	0.123	0.083	0.038	0.039
	(0.122)	(0.100)	(0.103)	(0.106)
Social Infrastructure		2.359***	2.389***	2.387***
		(0.269)	(0.257)	(0.266)
Ethnic Fractionalization			-0.684**	-0.738
			(0.277)	(0.781)
Ethnic Fractionalization Sqr.				0.064
				(0.923)
Optimal Diversity	0.713	0.725	0.724	0.724***
	(0.100)	(0.036)	(0.118)	(0.095)
Observations	109	109	109	109
R-squared	0.74	0.84	0.85	0.85

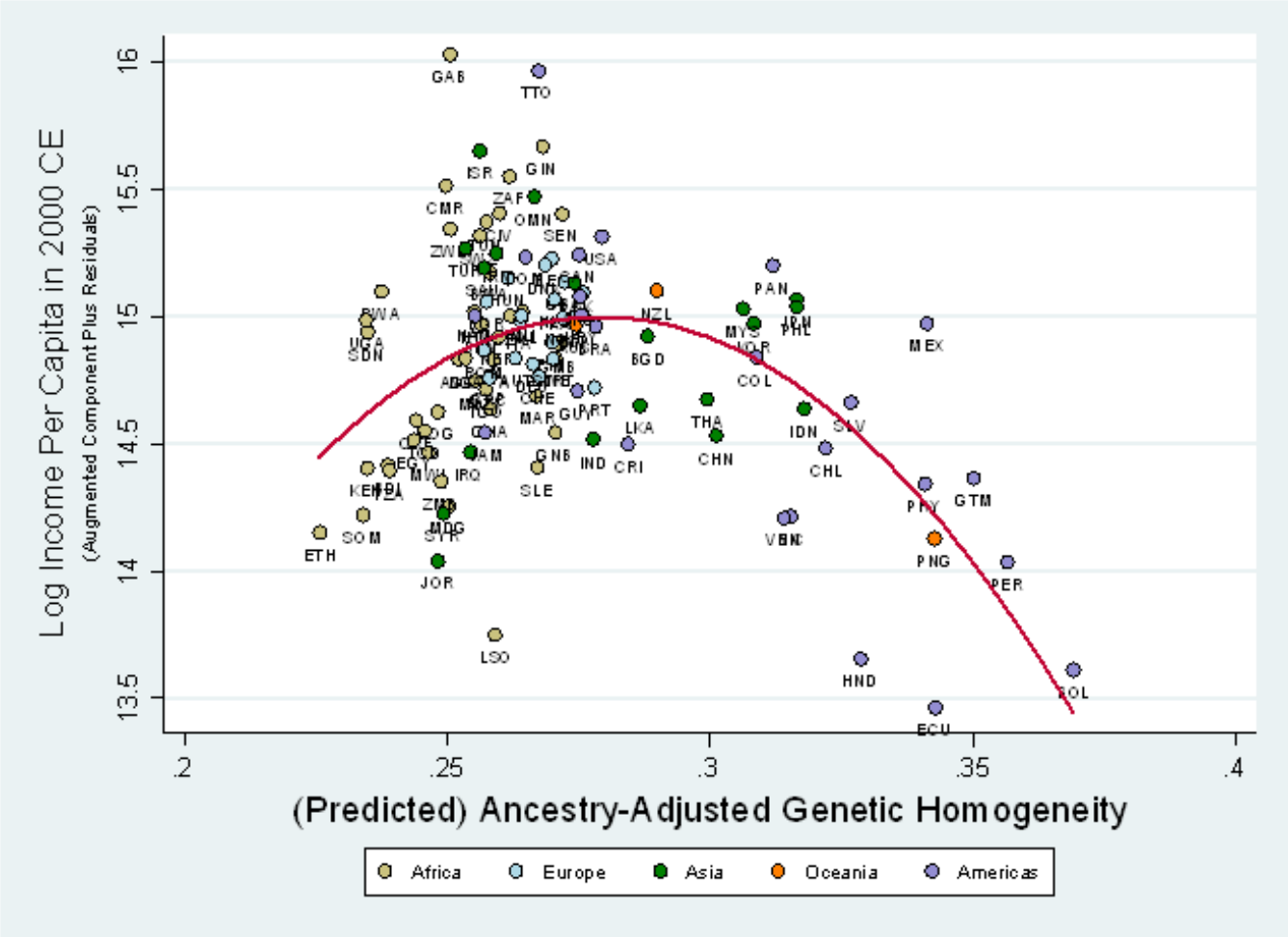
Notes: Bootstrap standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1;
All regressions include continent and SSA dummies.

Diversity and Other Determinants of Economic Development in 2000

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable is Log Income Per Capita in 2000 CE					
Adj. Diversity	315.282***	225.858***	204.102***	245.377***	277.342***	215.675***
	(84.215)	(67.669)	(66.984)	(66.845)	(70.232)	(63.954)
Adj. Diversity Sqr.	-220.980***	-155.826***	-140.850***	-170.036***	-192.386***	-150.871***
	(59.562)	(47.962)	(47.393)	(47.195)	(49.675)	(45.554)
Log Adj. Trans. Timing	-0.273	-0.092	-0.062	0.352	0.396*	-0.046
	(0.269)	(0.200)	(0.203)	(0.242)	(0.233)	(0.208)
Log Arable % of Land	-0.218***	-0.159***	-0.163***	-0.211***	-0.183***	-0.084
	(0.061)	(0.049)	(0.050)	(0.047)	(0.051)	(0.056)
Log Absolute Latitude	0.123	0.083	0.080	0.119	0.009	-0.006
	(0.122)	(0.100)	(0.101)	(0.097)	(0.108)	(0.087)
Social Infrastructure		2.359***	2.069***	2.072***	1.826***	0.880**
		(0.269)	(0.377)	(0.375)	(0.417)	(0.418)
Democracy			0.036			
			(0.029)			
Ethnic Fractionalization				-0.505	-0.333	-0.122
				(0.319)	(0.280)	(0.265)
% Population at Risk of Contracting Malaria					-0.502	-0.723**
					(0.351)	(0.353)
Avg. Years of Schooling						0.134***
						(0.042)
Optimal Diversity	0.713	0.725	0.725	0.722	0.721	0.715
	(0.014)	(0.032)	(0.045)	(0.014)	(0.008)	(0.073)
OPEC Dummy	No	No	No	No	Yes	Yes
Legal Origin Dummies	No	No	No	Yes	Yes	Yes
Major Religion Shares	No	No	No	Yes	Yes	Yes
Observations	109	109	109	109	109	94
R-squared	0.74	0.84	0.85	0.87	0.90	0.93

Notes: Bootstrap standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Diversity and Comparative Development in 2000



Interpretations - Comparative Development in 2000

- Optimal GD $= 0.7208 \approx$ GD in US $= 0.7206$
- Increasing the diversity of the most homogenous country (Bolivia; 0.6309) to that of the US $\Longrightarrow$ raise income per capita by a factor of 5.4 (increasing income in Bolivia from 9% to 40% of that of the US)
- Decreasing the diversity of the most heterogenous country (Ethiopia; 0.7743) to that of the US $\Longrightarrow$ raise income per capita by a factor of 1.7 (increasing income in Ethiopia from 2% to 4% of that of the US)

Interpretations - Comparative Development in 2000

- Optimal GD $= 0.7208 \approx$ GD in US $= 0.7206$
- Increasing the diversity of the most homogenous country (Bolivia; 0.6309) to that of the US $\Longrightarrow$ raise income per capita by a factor of 5.4 (increasing income in Bolivia from 9% to 40% of that of the US)
- Decreasing the diversity of the most heterogenous country (Ethiopia; 0.7743) to that of the US $\Longrightarrow$ raise income per capita by a factor of 1.7 (increasing income in Ethiopia from 2% to 4% of that of the US)

Interpretations - Comparative Development in 2000

- Optimal GD $= 0.7208 \approx$ GD in US $= 0.7206$
- Increasing the diversity of the most homogenous country (Bolivia; 0.6309) to that of the US $\Longrightarrow$ raise income per capita by a factor of 5.4 (increasing income in Bolivia from 9% to 40% of that of the US)
- Decreasing the diversity of the most heterogenous country (Ethiopia; 0.7743) to that of the US $\Longrightarrow$ raise income per capita by a factor of 1.7 (increasing income in Ethiopia from 2% to 4% of that of the US)

Interpretations - Comparative Development in 2000

- Increasing the diversity of the most homogenous country (Bolivia) by 1 percentage point would raise its income per capita in the year 2000 by 39%
- Decreasing the diversity of the most heterogenous country (Ethiopia) by 1 percentage point would raise its income per capita by 21%
- 1 percentage point change in diversity at the optimum level would lower income per capita by 1.9%

Addressing Endogenous Post-Columbian Migrations

	(1) Full Sample	(2) Non OECD	(3) w/o Neo Europes	(4) w/o Latin America	(5) w/o Sub Saharan	(6) >0.97 Indigenous
	Dependent Variable is Log Income Per Capita in 2000 CE					
Adj. Diversity	277.342***	271.979***	261.367***	412.222***	264.805**	304.735**
	(70.232)	(88.479)	(70.533)	(148.584)	(111.365)	(111.588)
Adj. Diversity Sqr.	-192.386***	-188.974***	-181.811***	-287.067***	-183.863**	-213.389**
	(49.675)	(62.096)	(49.671)	(101.906)	(80.398)	(77.255)
Log Adj. Trans. Timing	0.396*	0.390	0.355	0.518*	0.068	0.448*
	(0.233)	(0.281)	(0.231)	(0.298)	(0.442)	(0.254)
Log Arable % of Land	-0.183***	-0.236***	-0.201***	-0.189***	-0.211**	-0.104
	(0.051)	(0.060)	(0.055)	(0.050)	(0.097)	(0.061)
Log Absolute Latitude	0.009	-0.021	-0.025	-0.139	0.218	-0.074
	(0.108)	(0.119)	(0.111)	(0.126)	(0.242)	(0.130)
Social Infrastructure	1.826***	1.313**	1.416***	2.044***	1.585***	1.311*
	(0.417)	(0.579)	(0.507)	(0.545)	(0.486)	(0.716)
Ethnic Fractionalization	-0.333	-0.437	-0.390	-0.752**	0.104	-0.044
	(0.280)	(0.375)	(0.300)	(0.348)	(0.408)	(0.412)
% Population at Risk of Contracting Malaria	-0.502	-0.605	-0.591	-0.308	-0.425	-0.153
	(0.351)	(0.381)	(0.370)	(0.486)	(0.581)	(0.434)
% Population Living in Tropical Zones	-0.319	-0.196	-0.302	-0.520**	-0.528	-0.339
	(0.204)	(0.239)	(0.219)	(0.252)	(0.341)	(0.312)
Optimal Diversity	0.721	0.720	0.719	0.718	0.720	0.714
	(0.083)	(0.085)	(0.015)	(0.023)	(0.180)	(0.012)
Observations	109	83	105	87	71	37
R-squared	0.90	0.82	0.89	0.93	0.86	0.98

Notes: Bootstrap standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

The Costs of Genetic Diversity

	(1)	(2)	(3)	(4)
	Interpersonal Trust		# of Ethnic Wars 1946-2001	
Predicted Diversity	-1.959**	-2.200**	22.329***	22.341***
(Ancestry Adjusted)	(0.889)	(0.872)	(7.285)	(7.336)
Log Transition Timing	0.068	0.079	-0.603*	-0.601*
(Ancestry Adjusted)	(0.069)	(0.068)	(0.317)	(0.321)
Log % of Arable Land	0.002	-0.004	0.100	0.102
	(0.019)	(0.019)	(0.101)	(0.102)
Log Absolute Latitude	-0.003	0.017	-0.256	-0.246
	(0.030)	(0.029)	(0.230)	(0.242)
Social Infrastructure	0.191**	0.108	0.072	0.025
	(0.078)	(0.088)	(0.439)	(0.493)
Ethnic Fractionalization	0.062	0.053	-0.743	-0.726
	(0.077)	(0.072)	(0.702)	(0.705)
% of Population of European	0.051	0.005	-0.339	-0.516
Descent	(0.106)	(0.101)	(1.139)	(1.207)
% of Population at Risk of	-0.061	-0.022	0.361	0.370
Contracting Malaria	(0.094)	(0.089)	(0.549)	(0.554)
Terrain Ruggedness	-0.240	-0.193	1.261**	1.305*
	(0.148)	(0.137)	(0.619)	(0.670)
% of Population Living in	-0.067	0.003	0.349	0.362
Tropical Zones	(0.060)	(0.066)	(0.660)	(0.687)
Mean Distance to Nearest	0.091	0.114*	1.063**	1.077*
Waterway	(0.063)	(0.061)	(0.525)	(0.545)
Years of Schooling		0.021*		0.016
		(0.012)		(0.073)
Observations	58	58	69	69
R-squared	0.79	0.80	0.59	0.59

The Benefits of Genetic Diversity

	(1)	(2)	(3)	(4)
	Patent Applications 1980-2000		Scientific Articles 1980-2000	
Predicted Diversity	0.851**	0.673**	2.290***	1.816***
(Ancestry Adjusted)	(0.343)	(0.312)	(0.576)	(0.541)
Log Transition Timing	-0.021	-0.007	-0.091*	-0.076
(Ancestry Adjusted)	(0.041)	(0.042)	(0.048)	(0.048)
Log % of Arable Land	-0.003	-0.000	0.009	0.007
	(0.013)	(0.013)	(0.016)	(0.015)
Log Absolute Latitude	0.010	0.017	0.045*	0.055**
	(0.015)	(0.013)	(0.024)	(0.023)
Social Infrastructure	0.241***	0.177**	0.685***	0.548***
	(0.078)	(0.072)	(0.117)	(0.127)
Ethnic Fractionalization	0.003	-0.008	0.095	0.073
	(0.059)	(0.060)	(0.096)	(0.096)
% of Population of European	0.029	-0.042	0.042	-0.040
Descent	(0.067)	(0.062)	(0.090)	(0.081)
% of Population at Risk of	0.031	0.035	0.102*	0.131***
Contracting Malaria	(0.049)	(0.043)	(0.055)	(0.048)
Terrain Ruggedness	-0.086	-0.060	-0.349*	-0.269
	(0.098)	(0.095)	(0.177)	(0.169)
% of Population Living in	-0.021	0.004	0.018	0.049
Tropical Zones	(0.036)	(0.031)	(0.058)	(0.055)
Mean Distance to Nearest	-0.037	-0.031	0.105***	0.118***
Waterway	(0.044)	(0.044)	(0.038)	(0.035)
Years of Schooling		0.020***		0.032***
		(0.007)		(0.008)
Observations	77	77	93	93
R-squared	0.74	0.77	0.80	0.82

Concluding Remarks

- Deep-rooted factors, determined tens of thousands of years ago, had a significant effect on the:
  - course of economic development from the dawn of human civilization to the contemporary era
  - contemporary differences in economic development across countries
- Variation in migratory distance from the cradle of humankind in East Africa to settlements across the globe
  - $\Longrightarrow$ genetic diversity $\Longrightarrow$ comparative development
  - The optimal level of diversity has increased in the post-industrialization era.

Concluding Remarks

- Deep-rooted factors, determined tens of thousands of years ago, had a significant effect on the:
  - course of economic development from the dawn of human civilization to the contemporary era
  - contemporary differences in economic development across countries
- Variation in migratory distance from the cradle of humankind in East Africa to settlements across the globe
  - $\Longrightarrow$ genetic diversity $\Longrightarrow$ comparative development
  - The optimal level of diversity has increased in the post-industrialization era.

Concluding Remarks

- Deep-rooted factors, determined tens of thousands of years ago, had a significant effect on the:
  - course of economic development from the dawn of human civilization to the contemporary era
  - contemporary differences in economic development across countries
- Variation in migratory distance from the cradle of humankind in East Africa to settlements across the globe
  - $\Longrightarrow$ genetic diversity $\Longrightarrow$ comparative development
  - The optimal level of diversity has increased in the post-industrialization era.

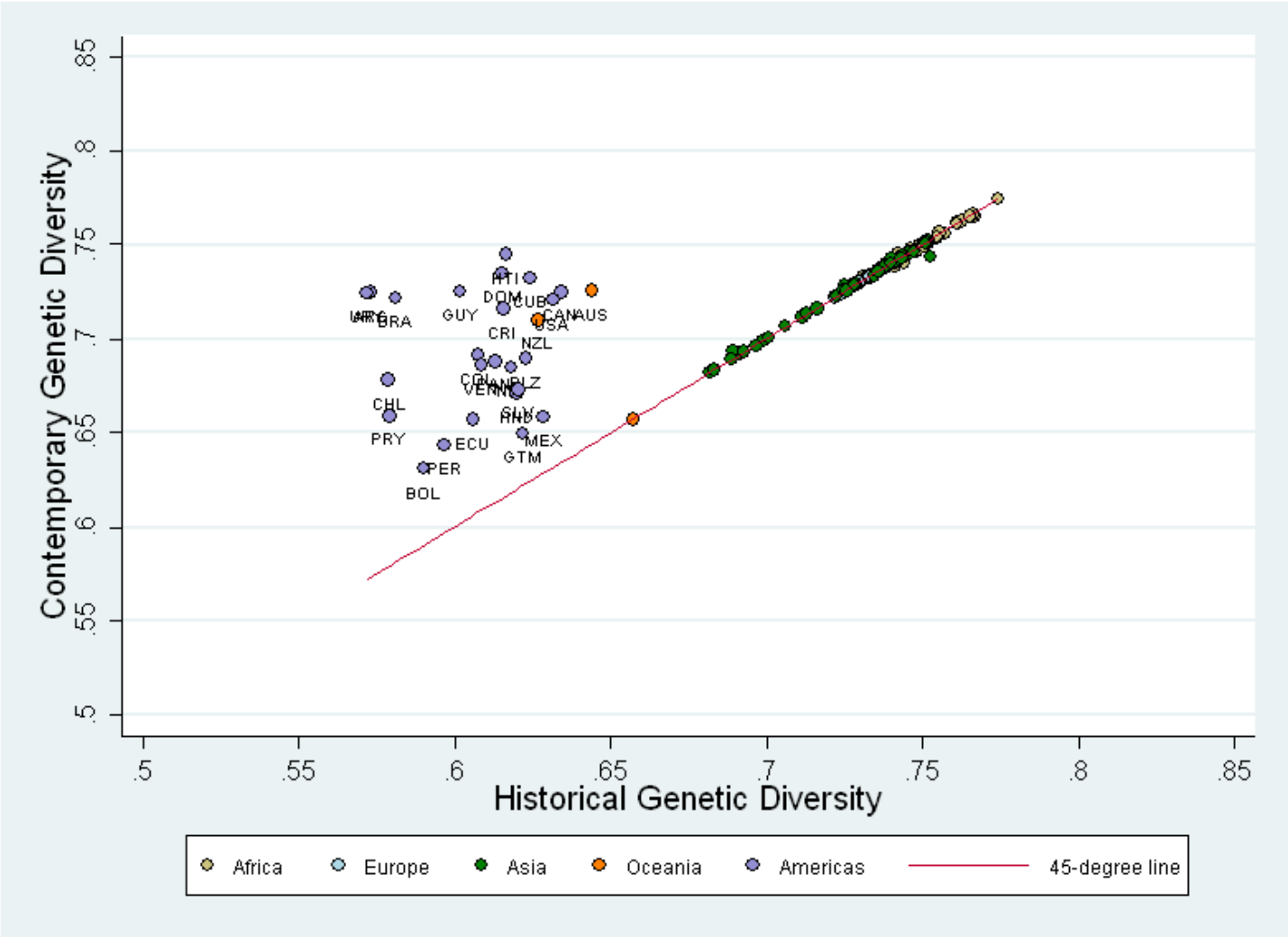
Concluding Remarks

- Deep-rooted factors, determined tens of thousands of years ago, had a significant effect on the:
  - course of economic development from the dawn of human civilization to the contemporary era
  - contemporary differences in economic development across countries
- Variation in migratory distance from the cradle of humankind in East Africa to settlements across the globe
  - $\Longrightarrow$ genetic diversity $\Longrightarrow$ comparative development
  - The optimal level of diversity has increased in the post-industrialization era.

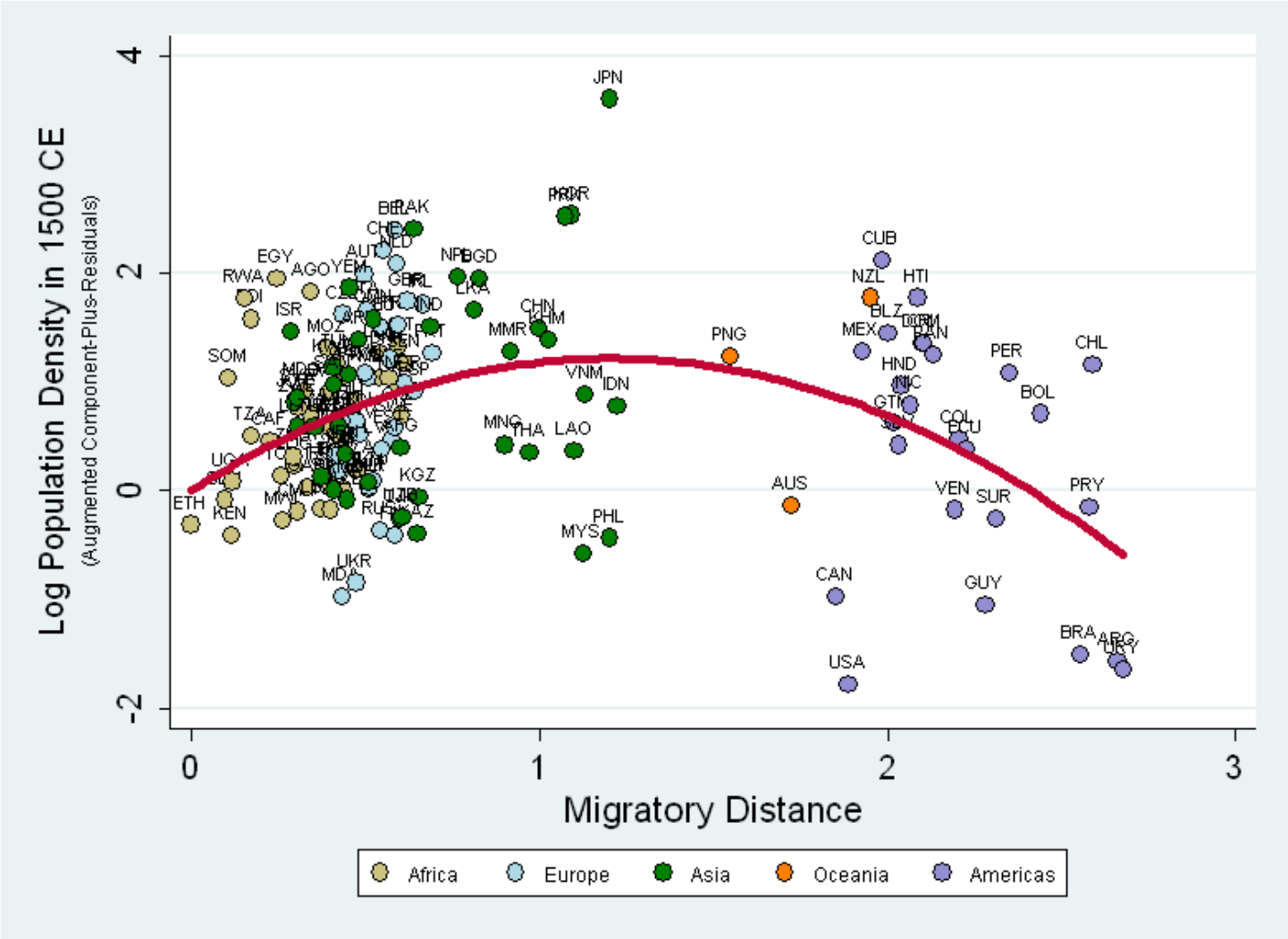
Concluding Remarks

- Deep-rooted factors, determined tens of thousands of years ago, had a significant effect on the:
  - course of economic development from the dawn of human civilization to the contemporary era
  - contemporary differences in economic development across countries
- Variation in migratory distance from the cradle of humankind in East Africa to settlements across the globe
  - $\Longrightarrow$ genetic diversity $\Longrightarrow$ comparative development
  - The optimal level of diversity has increased in the post-industrialization era.

Genetic Diversity in 1500 vs 2000



Migratory Distance and Pre-Colonial Comparative Development



Diversity and Comparative Development in 2000

Adjusted versus Unadjusted Diversity

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable is Log Income Per Capita in 2000 CE					
Adj. Diversity	556.439***	254.906***			533.983***	387.314**
	(134.034)	(89.814)			(167.245)	(188.496)
Adj. Diversity Sqr.	-397.224***	-176.907***			-377.365***	-273.925**
	(93.793)	(63.553)			(119.747)	(136.677)
Unadj. Diversity			140.903***	10.152	1.670	-64.226
			(53.395)	(54.148)	(70.175)	(82.555)
Unadj. Diversity Sqr.			-107.686***	-7.418	-4.057	51.016
			(39.408)	(39.003)	(53.740)	(65.051)
Continent Dummies	No	Yes	No	Yes	No	Yes
Observations	143	143	143	143	143	143
R-squared	0.13	0.47	0.08	0.45	0.14	0.48
P-value for:						
Joint Sig. of Adjusted Diversity and its Sqr.					0.009	0.038
Joint Sig. of Unadjusted Diversity and its Sqr.					0.399	0.741

Notes: Bootstrap standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Migratory Distance and Pre-Colonial Comparative Development

Migratory Distance from East Africa and Pre-Colonial Comparative Development

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable is Log Population Density in:					
	1500 CE				1000 CE	1 CE
Mig. Distance	1.792**	1.663**	1.379**	2.011***	1.865**	2.017**
	(0.794)	(0.728)	(0.606)	(0.670)	(0.768)	(0.867)
Mig. Distance Sqr.	-1.012***	-0.867***	-0.787***	-0.834***	-0.832**	-0.952**
	(0.280)	(0.263)	(0.231)	(0.314)	(0.378)	(0.450)
Log Transition Timing		1.047***	1.160***	1.235***	1.603***	2.127***
		(0.189)	(0.144)	(0.224)	(0.246)	(0.357)
Log Absolute Latitude			-0.342***	-0.417***	-0.373***	-0.115
			(0.091)	(0.119)	(0.128)	(0.127)
Log Arable % of Land			0.401***	0.393***	0.370***	0.348***
			(0.091)	(0.097)	(0.108)	(0.121)
Log Agri. Suitability			0.305***	0.257***	0.190*	0.210*
			(0.091)	(0.094)	(0.102)	(0.117)
Optimal Mig. Distance	0.886***	0.958***	0.876***	1.206***	1.121***	1.060***
	(0.156)	(0.142)	(0.144)	(0.276)	(0.309)	(0.260)
Continent Dummies	No	No	No	Yes	Yes	Yes
Observations	145	145	145	145	140	126
R-squared	0.22	0.38	0.67	0.69	0.62	0.61

Notes: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Migratory Distance and Pre-Colonial Comparative Development

Robustness to Alternative Distances

	(1)	(2)	(3)	(4)	(5)
	Dependent Variable is: Log Population Density in 1500 CE				
	Distances from:				
	Addis Ababa	Addis Ababa	London	Tokyo	Mexico City
Mig. Distance	1.379**		-0.040	0.052	-0.063
	(0.606)		(0.063)	(0.145)	(0.099)
Mig. Distance Sqr.	-0.787***		-0.002	-0.006	0.005
	(0.231)		(0.002)	(0.007)	(0.004)
Aerial Distance		-0.008			
		(0.106)			
Aerial Distance Sqr.		-0.005			
		(0.006)			
Log Transition Timing	1.160***	1.158***	1.003***	1.047***	1.619***
	(0.144)	(0.138)	(0.164)	(0.225)	(0.277)
Log Absolute Latitude	-0.342***	-0.263***	-0.358***	-0.334***	-0.239***
	(0.091)	(0.097)	(0.112)	(0.099)	(0.083)
Log Arable % of Land	0.401***	0.488***	0.357***	0.532***	0.493***
	(0.091)	(0.102)	(0.092)	(0.089)	(0.094)
Log Agri. Suitability	0.305***	0.254**	0.344***	0.178**	0.261***
	(0.091)	(0.102)	(0.092)	(0.080)	(0.092)
Observations	145	145	145	145	145
R-squared	0.67	0.59	0.67	0.59	0.63

Notes: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Migratory Distance and Contemporary Comparative Development

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable is Log Income Per Capita in 2000 CE					
Adj. Mig. Distance	2.325***	2.034***	1.167**	1.355**	1.555***	1.357***
	(0.620)	(0.668)	(0.560)	(0.528)	(0.499)	(0.477)
Adj. Mig. Distance Sqr.	-1.354***	-1.260***	-0.889***	-0.970***	-1.097***	-0.860***
	(0.293)	(0.307)	(0.261)	(0.248)	(0.231)	(0.212)
Log Adj. Transition Timing		-0.273	-0.092	0.352	0.396*	-0.046
		(0.261)	(0.193)	(0.222)	(0.217)	(0.181)
Log Arable % of Land		-0.218***	-0.159***	-0.211***	-0.183***	-0.084*
		(0.057)	(0.046)	(0.044)	(0.045)	(0.046)
Log Absolute Latitude		0.123	0.083	0.119	0.009	-0.006
		(0.114)	(0.097)	(0.094)	(0.093)	(0.072)
Social Infrastructure			2.359***	2.072***	1.826***	0.880**
			(0.244)	(0.342)	(0.364)	(0.357)
Ethnic Fractionalization				-0.505*	-0.333	-0.122
				(0.297)	(0.254)	(0.225)
% Population at Risk of Contracting Malaria					-0.502	-0.723**
					(0.312)	(0.305)
Waterway					(0.173)	(0.145)
Avg. Years of Schooling						0.134***
						(0.032)
Optimal Adj. Mig. Distance	0.859***	0.807***	0.657***	0.699***	0.709***	0.788***
	(0.087)	(0.107)	(0.147)	(0.122)	(0.101)	(0.108)
OPEC Dummy	No	No	No	No	Yes	Yes
Legal Origin Dummies	No	No	No	Yes	Yes	Yes
Major Religion Shares	No	No	No	Yes	Yes	Yes
Observations	109	109	109	109	109	94
R-squared	0.69	0.74	0.84	0.87	0.90	0.93

Notes: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Migratory Distance and Contemporary Comparative Development

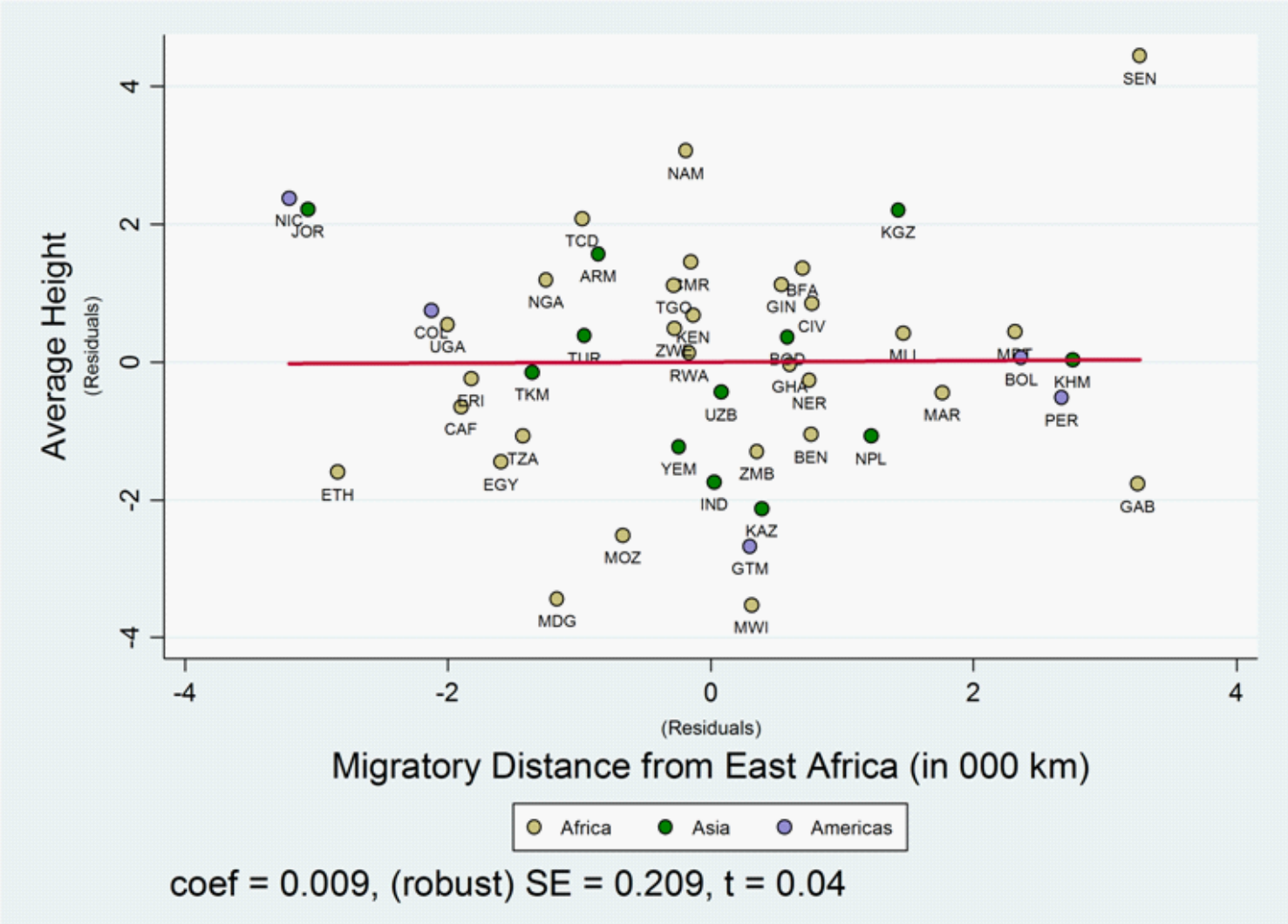
Robustness to Alternative Distances for Contemporary Development

	(1)	(2)	(3)
	Dependent Variable is Log Income Per Capita in 2000 CE		
Adj. Migratory Distance	4.992***	5.316***	5.643**
	(1.240)	(1.599)	(2.351)
Adj. Migratory Distance Sqr.	-2.610***	-2.745***	-2.879***
	(0.556)	(0.757)	(0.978)
Migratory Distance	0.781		
	(0.836)		
Migratory Distance Sqr.	-0.219		
	(0.257)		
Aerial Distance		0.645	
		(2.012)	
Aerial Distance Sqr.		-0.185	
		(1.130)	
Adj. Aerial Distance			0.425
			(3.304)
Adj. Aerial Distance Sqr.			-0.069
			(1.786)
Continent Dummies	No	No	No
Observations	109	109	109
R-squared	0.29	0.29	0.29

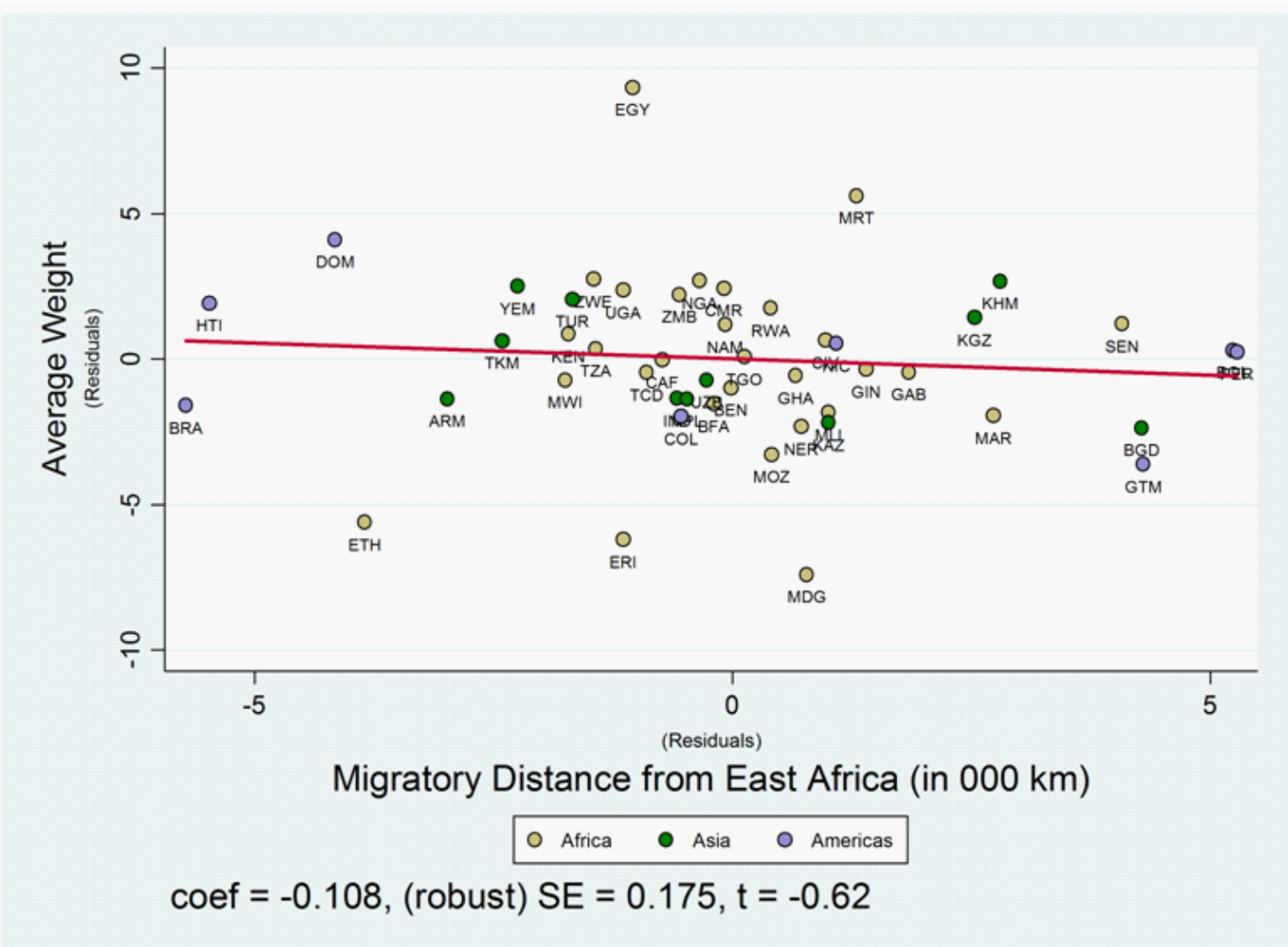
Notes: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Migratory Distance and Comparative Development

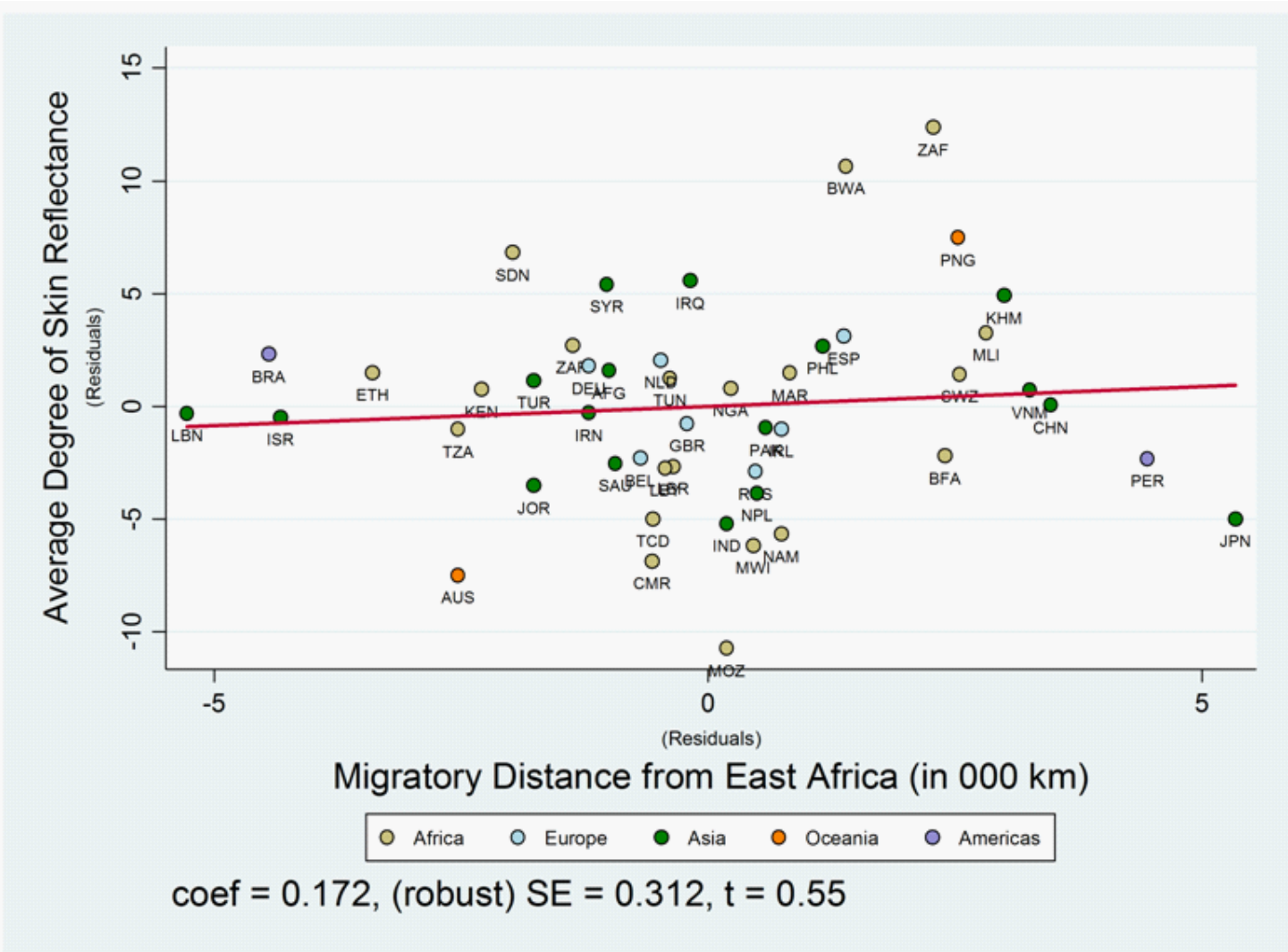
Orthogonality of Migratory Distance from Africa to Level Effects



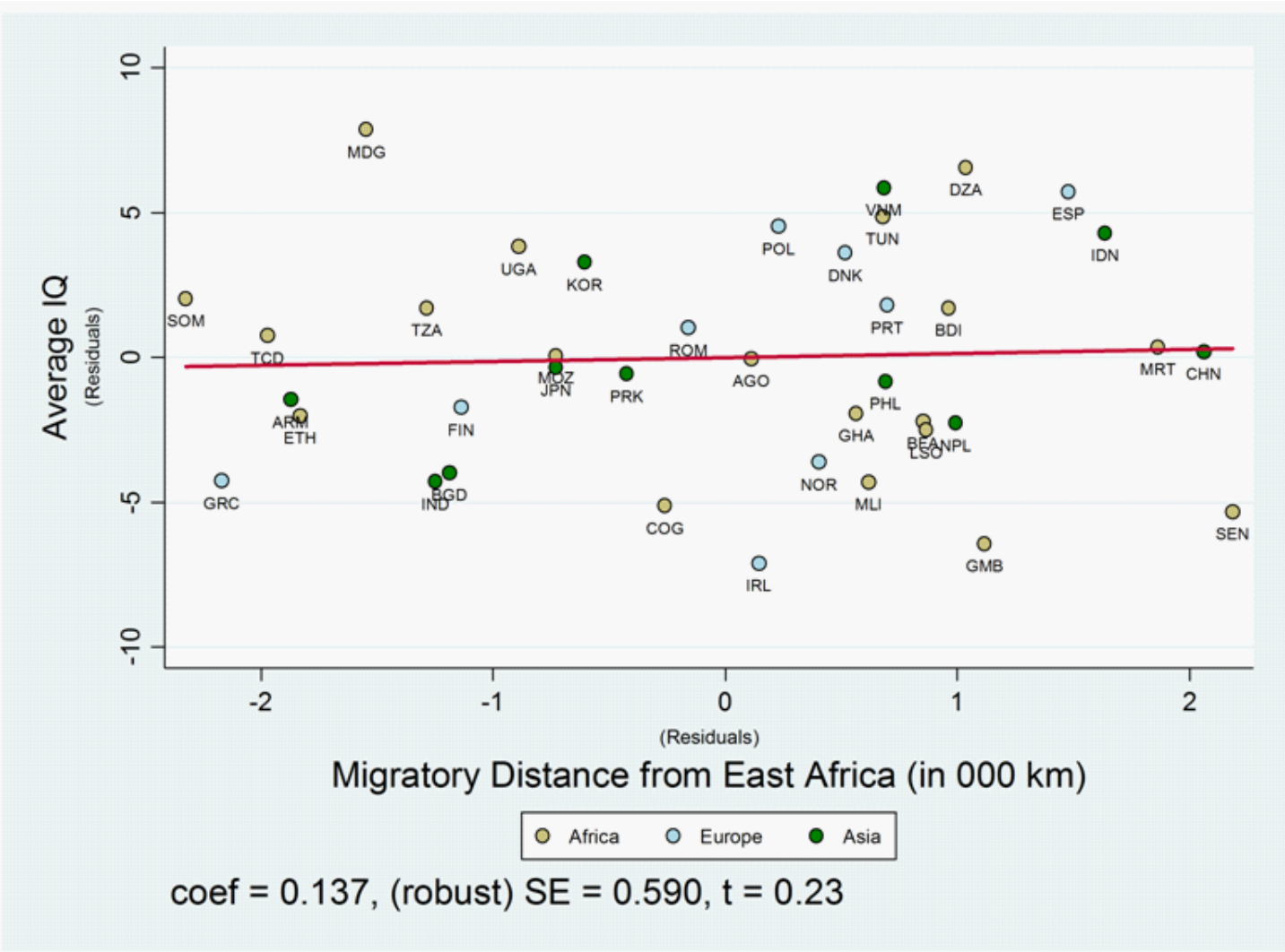
Orthogonality of Migratory Distance from Africa to Level Effects



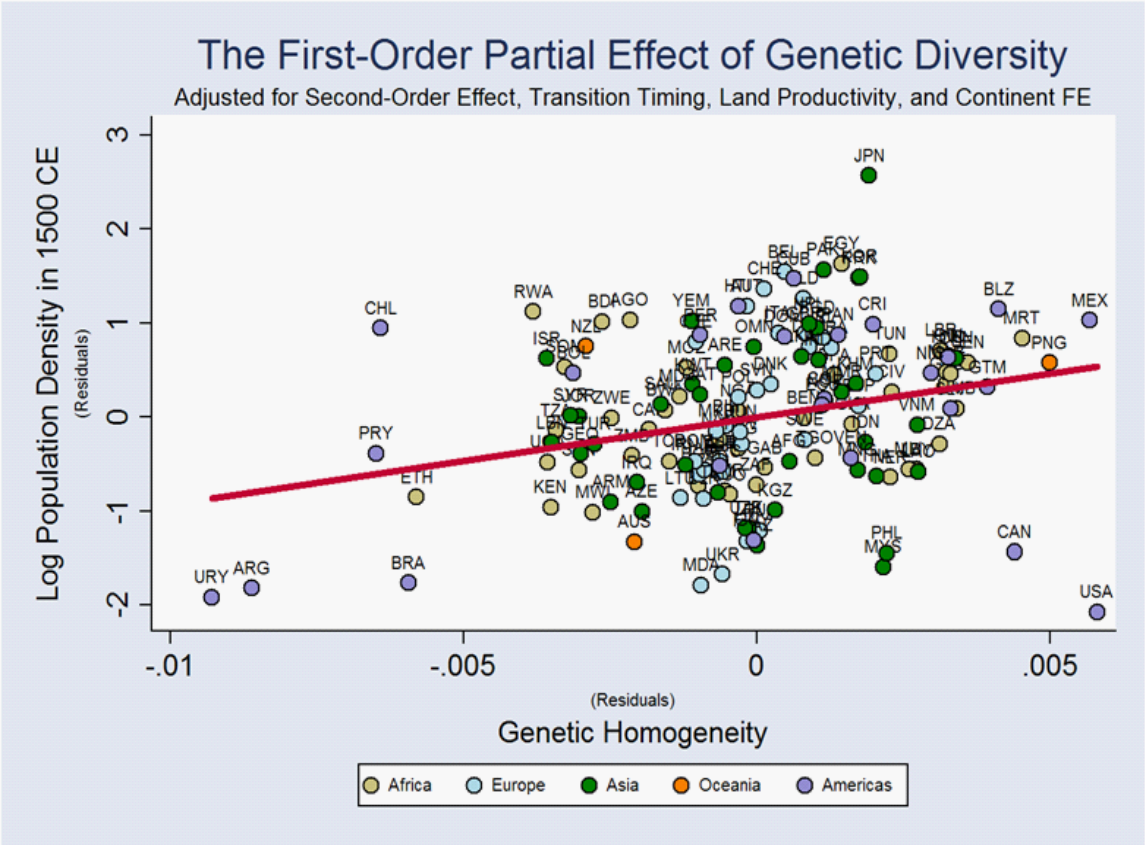
Orthogonality of Migratory Distance from Africa to Level Effects



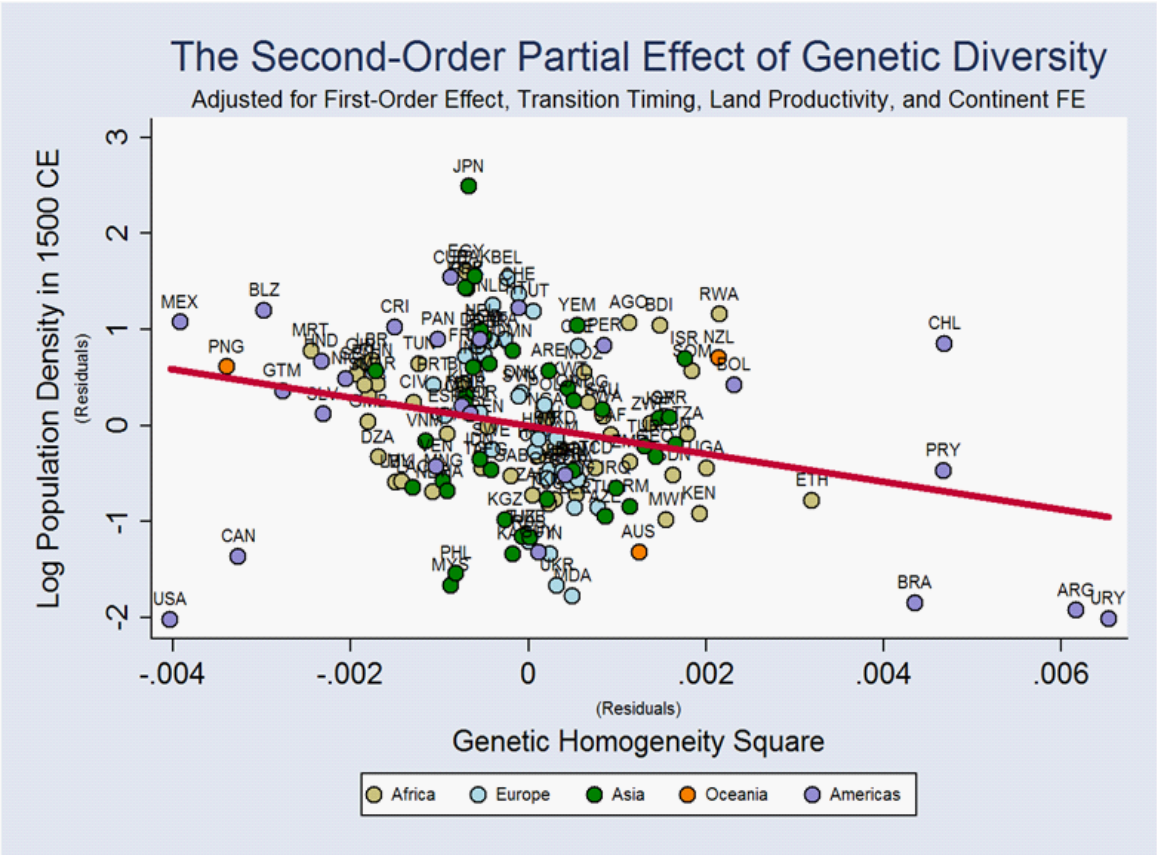
Orthogonality of Migratory Distance from Africa to Level Effects



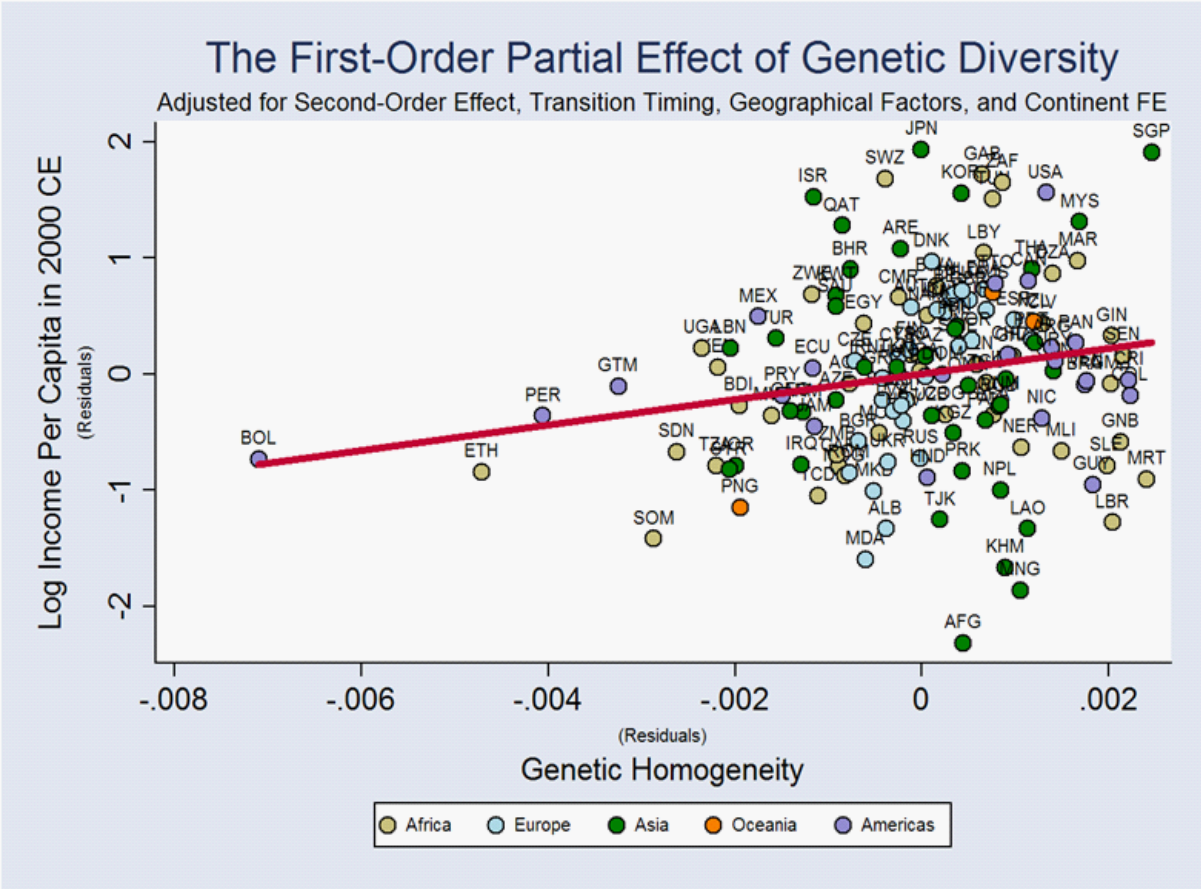
Diversity and Comparative Development in 1500 CE



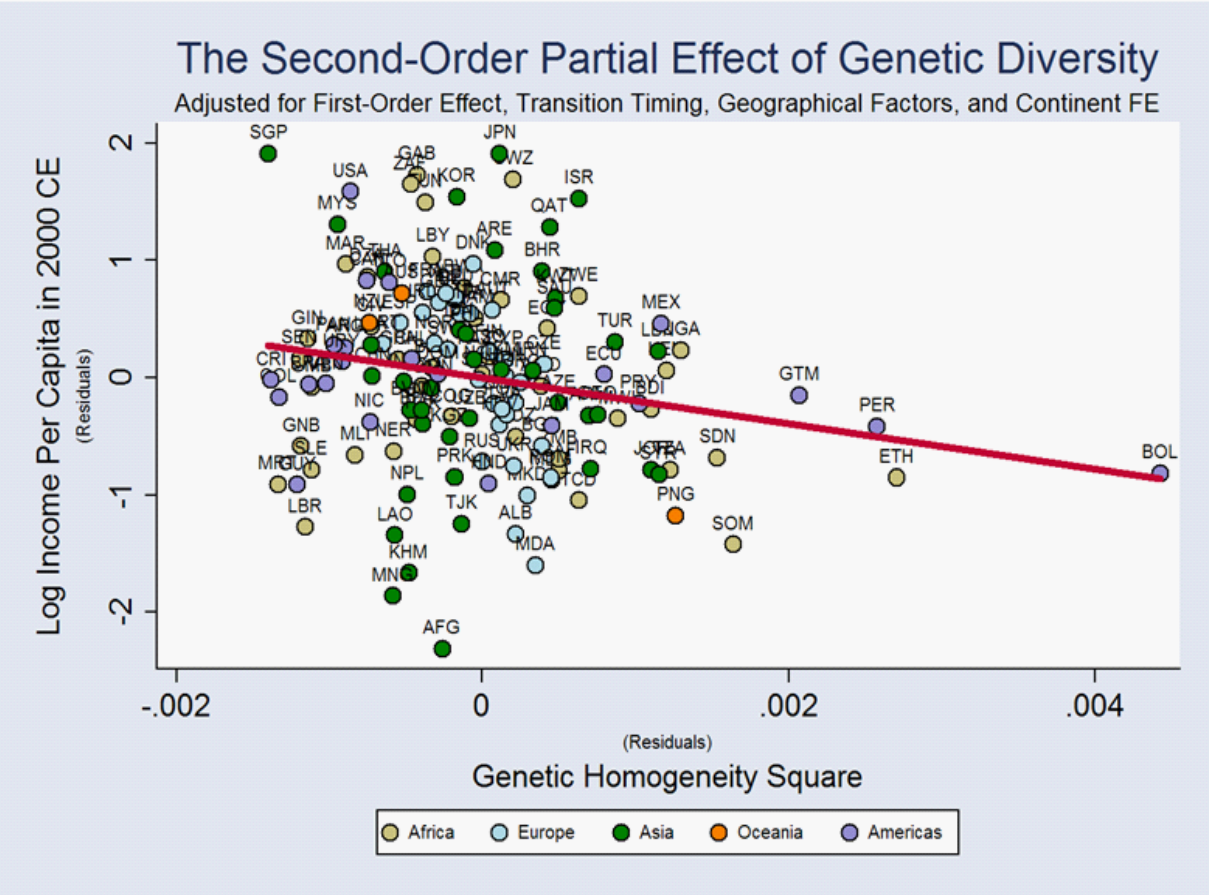
Diversity and Comparative Development in 1500 CE



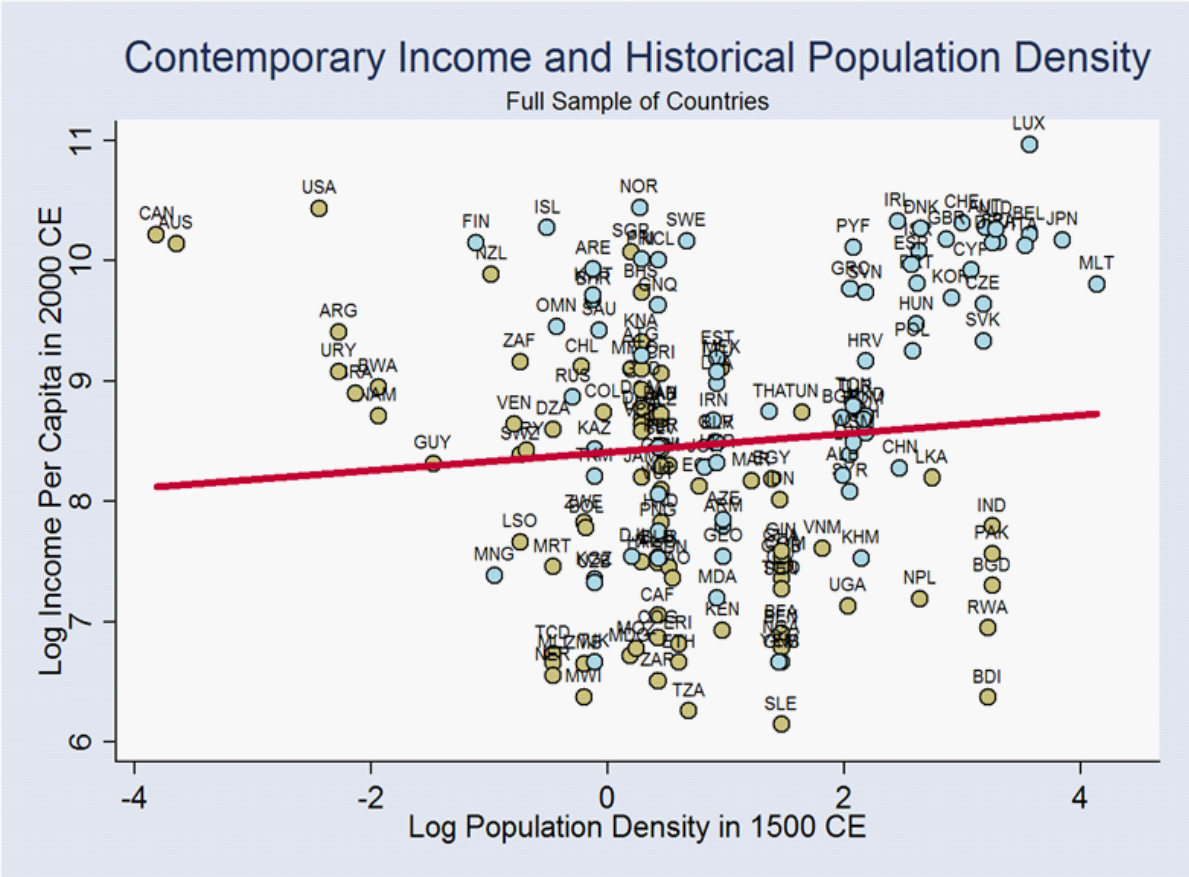
Diversity and Comparative Development in 1500 CE



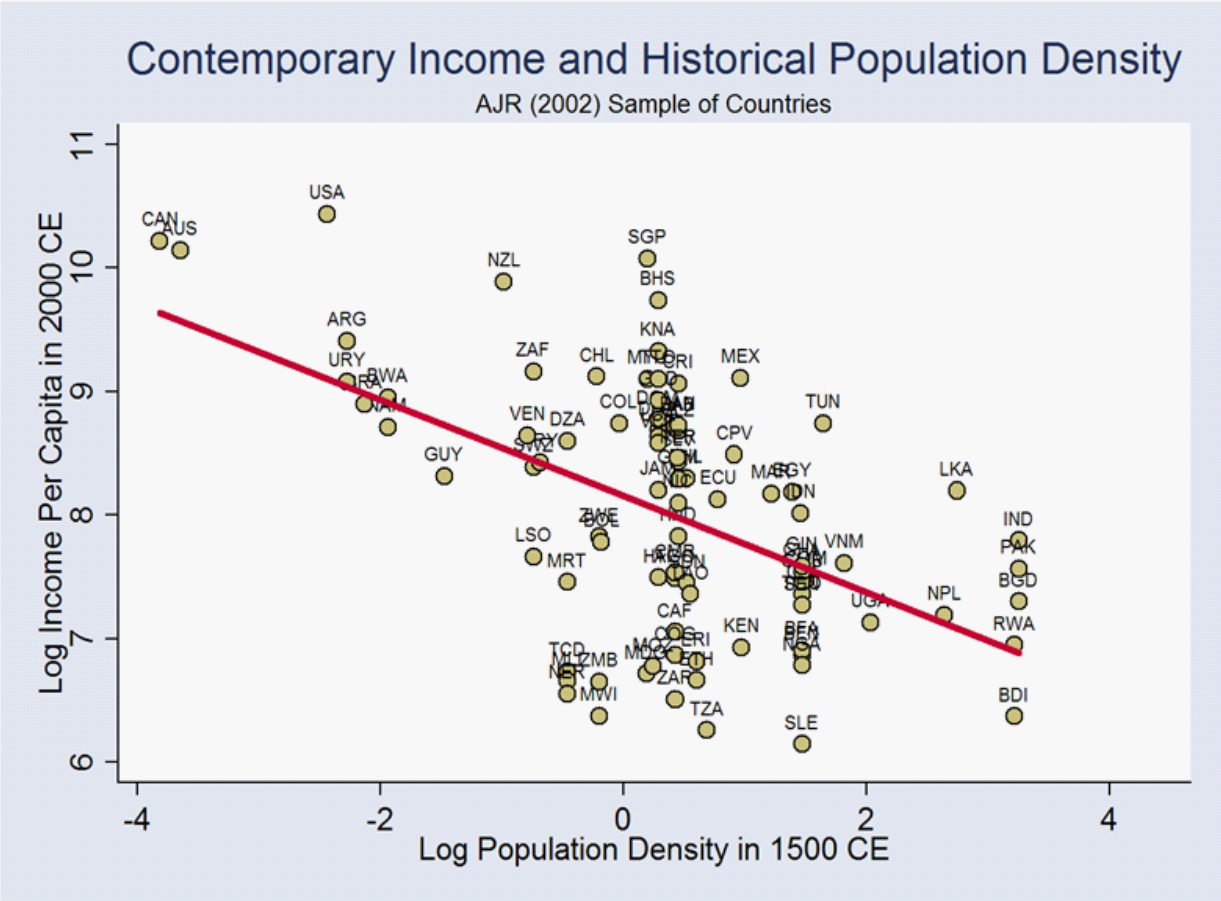
Diversity and Comparative Development in 1500 CE



Population Density in 1500 CE and Income Per Capita in 2000: Full Sample



Population Density in 1500 CE and Income Per Capita in 2000: AJR Sample



Data Sources

- Data on population density for countries (as defined by their modern national borders) in the years 1500CE, 1000CE and 1 CE is obtained from McEvedy and Jones (1978)
- Expected heterozygosity data on 53 ethnic groups is taken from the Human Genome Diversity Project (HGDP) and the genetic diversity study conducted by Ramachandran et al. (2005)
- Data on the timing of the transition to agriculture (in terms of years since the Neolithic Revolution) is provided by Olsson and Hibbs (2005) and Putterman (2007)
- Arable percentages and the suitability of land for agriculture come from the World Development Indicators and geological studies while latitudes are obtained from the CIA Factbook

The Human Genome Diversity Project (HGDP)

- Characteristics of the collection:
  - 1064 unrelated individuals from 53 ethnic groups worldwide
  - DNA extracted from B-lymphocytes, a class of white blood cells amenable to preservation and replication in the laboratory
  - Ethnic groups were chosen based on the criterion that they were in place prior to the great diasporas of the 16th century
  - Greater representation of countries (e.g., Pakistan and China) that have the necessary infrastructure for collecting DNA samples from different ethnic groups and making them available for research

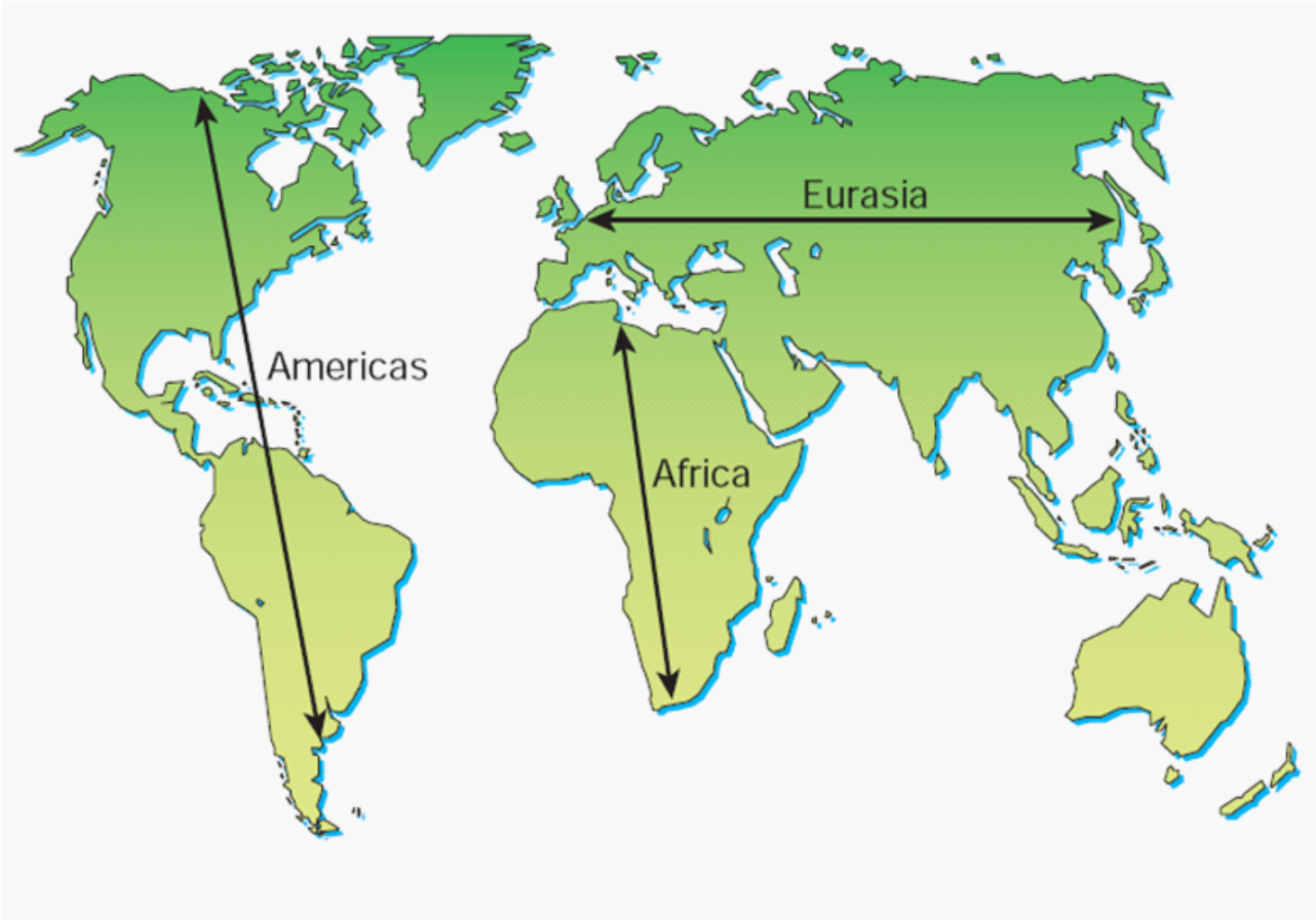
The Geographic Channel

- Population density should be positively related to the natural productivity of land
- Land productivity is proxied by:
  - Percentage of arable land
  - Suitability of land for agriculture
  - Absolute latitude
- Contemporary data: (Identifying assumption: The contemporary distribution of land quality does not differ significantly from that which existed during the pre-Colonial Malthusian era)

The Neolithic Channel (Diamond)

- The transition from hunting and gathering to agriculture permitted societies to generate resource surpluses that led to the emergence of a non-food-producing social class whose members were crucial for the development of new technologies
- The key factor that enabled the agricultural transition in some regions was the exogenous availability of certain wild species of plants and animals amenable for domestication, governed by:
  - Geographical factors that maximized the biodiversity in human habitats (climate, latitude, size of landmass)
  - Orientation of continents – played a crucial role in the diffusion of agricultural practices along similar latitudes

The Orientation of the Continents



Migratory Distance vs. Genetic Diversity

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	OLS	OLS	2SLS	2SLS
	Dependent Variable is Log Population Density in 1500 CE					
Genetic Diversity		255.220**	225.441***	203.815*	233.758***	181.932**
		(100.586)	(73.781)	(97.637)	(86.883)	(71.934)
Genetic Diversity Sqr.		-209.808**	-161.159**	-145.717*	-167.564**	-130.762**
		(73.814)	(56.155)	(80.414)	(65.729)	(59.269)
Mig. Distance	0.505***	0.070				
	(0.148)	(0.184)				
Mig. Distance Sqr.	-0.023***	-0.014				
	(0.006)	(0.009)				
Log Transition Timing			1.214***	1.135	1.183***	1.166**
			(0.373)	(0.658)	(0.338)	(0.475)
Log Arable % of Land			0.516***	0.545*	0.531***	0.545**
			(0.165)	(0.262)	(0.170)	(0.219)
Log Absolute Latitude			-0.162	-0.129	-0.169	-0.118
			(0.130)	(0.174)	(0.106)	(0.128)
Log Agri. Suitability			0.571*	0.587	0.558**	0.595**
			(0.294)	(0.328)	(0.256)	(0.256)
Optimal Diversity			0.699	0.699	0.698	0.696
			(0.015)	(0.055)	(0.015)	(0.045)
Continent Dummies	No	No	No	Yes	No	Yes
Observations	21	21	21	21	21	21
R-squared	0.34	0.46	0.89	0.90	–	–
Joint Sig. of Diversity and its Sqr.		0.023				
Joint Sig. of Distance and its Sqr.		0.235				
Overidentifying Restrictions					0.889	0.861
Exogeneity of Distance and its Sqr.					0.952	0.804

Related Literature

- This study is singular in its emphasis on:
  - Migratory distance "out of Africa"
  - genetic diversity *within* populations
- Earlier studies used genetic distance *between* populations – i.e., their mean genealogical unrelatedness
  - Spolaore and Wacziarg (2006): Genetic distance proxies for the effect of cultural differences on pairwise differences in income per capita (relative to the U.S.)

    $\Longrightarrow$ Interdependence across societies (through technological diffusion or trade) is a necessary condition for the effect of genetic differences on economic outcomes

Fundamental Differences with Earlier Studies

- Genetic diversity within a society has a significant impact, independently of society's position in the world economy

  Earlier: Interdependence across societies is a necessary condition for the effect of genetics on economic outcomes

- No hierarchy of genetic traits

  Earlier: Genetic traits more similar to those in the technological frontier (US) are beneficial

- Genetic diversity generates costs and benefits (hump-shaped relationship between diversity and development)

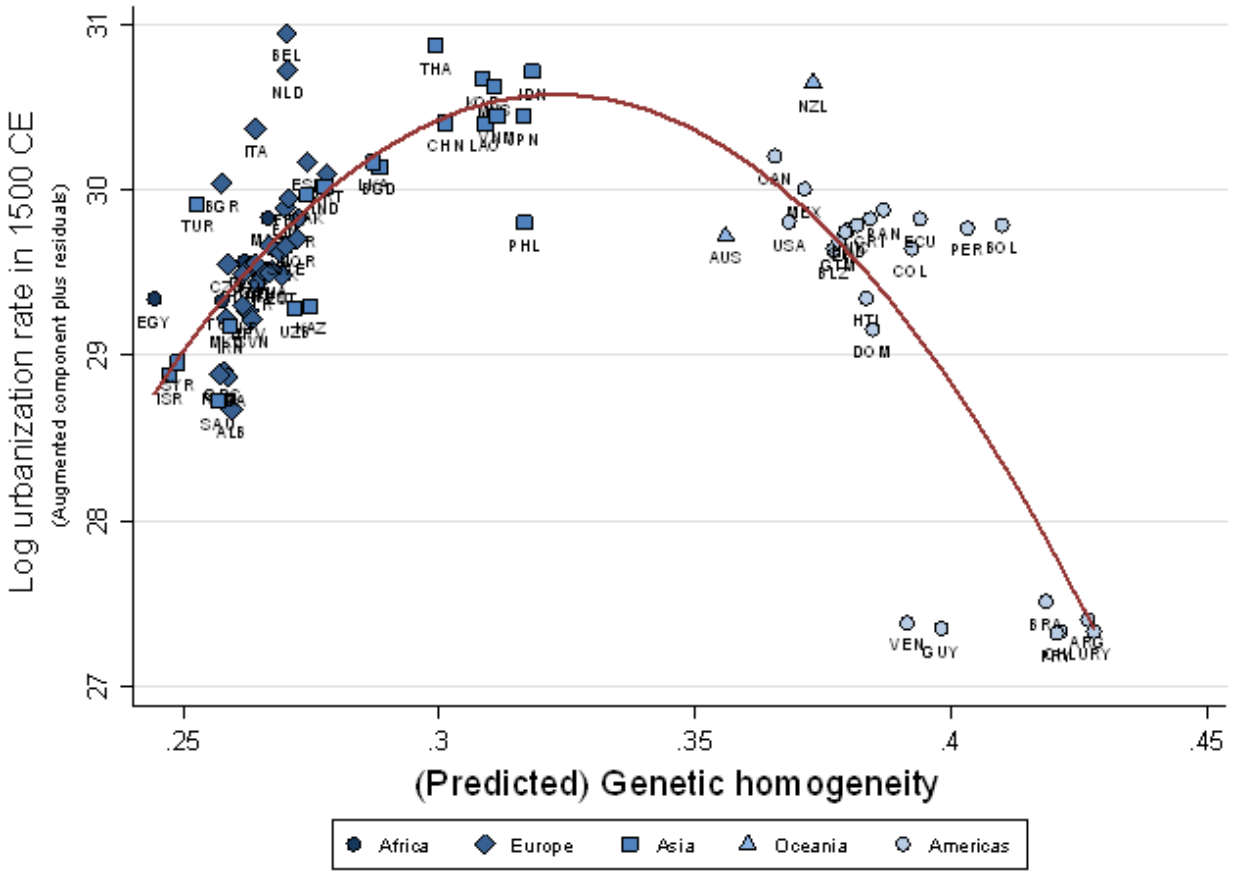
  Earlier: captured only the detrimental effects of genetic distance between societies

Diversity and Urbanization Rate in 1500 CE

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	OLS	OLS	OLS	OLS
	Dependent Variable is Log Urbanization Rate in 1500 CE					
Predicted diversity	120.583**		165.167***	93.467*	148.757***	234.410***
	(51.618)		(50.088)	(48.769)	(48.373)	(67.321)
Predicted diversity Square	-84.760**		-120.124***	-62.408*	-106.165***	-166.786***
	(38.423)		(37.208)	(36.650)	(36.506)	(48.780)
Log Neolithic transition timing		0.614***	0.457**		0.402**	0.752***
		(0.133)	(0.224)		(0.202)	(0.257)
Log arable land				-0.097**	-0.116***	-0.119**
				(0.043)	(0.044)	(0.052)
Log absolute latitude				-0.334**	-0.236	-0.151
				(0.151)	(0.155)	(0.170)
Log land suitability for agriculture				0.002	-0.036	0.031
				(0.057)	(0.058)	(0.059)
Continent fixed effects	No	No	No	No	No	Yes
Observations	80	80	80	80	80	80
R^2	0.30	0.17	0.35	0.40	0.44	0.51

Notes: Bootstrapped standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Diversity and Urbanization Rate in 1500 CE

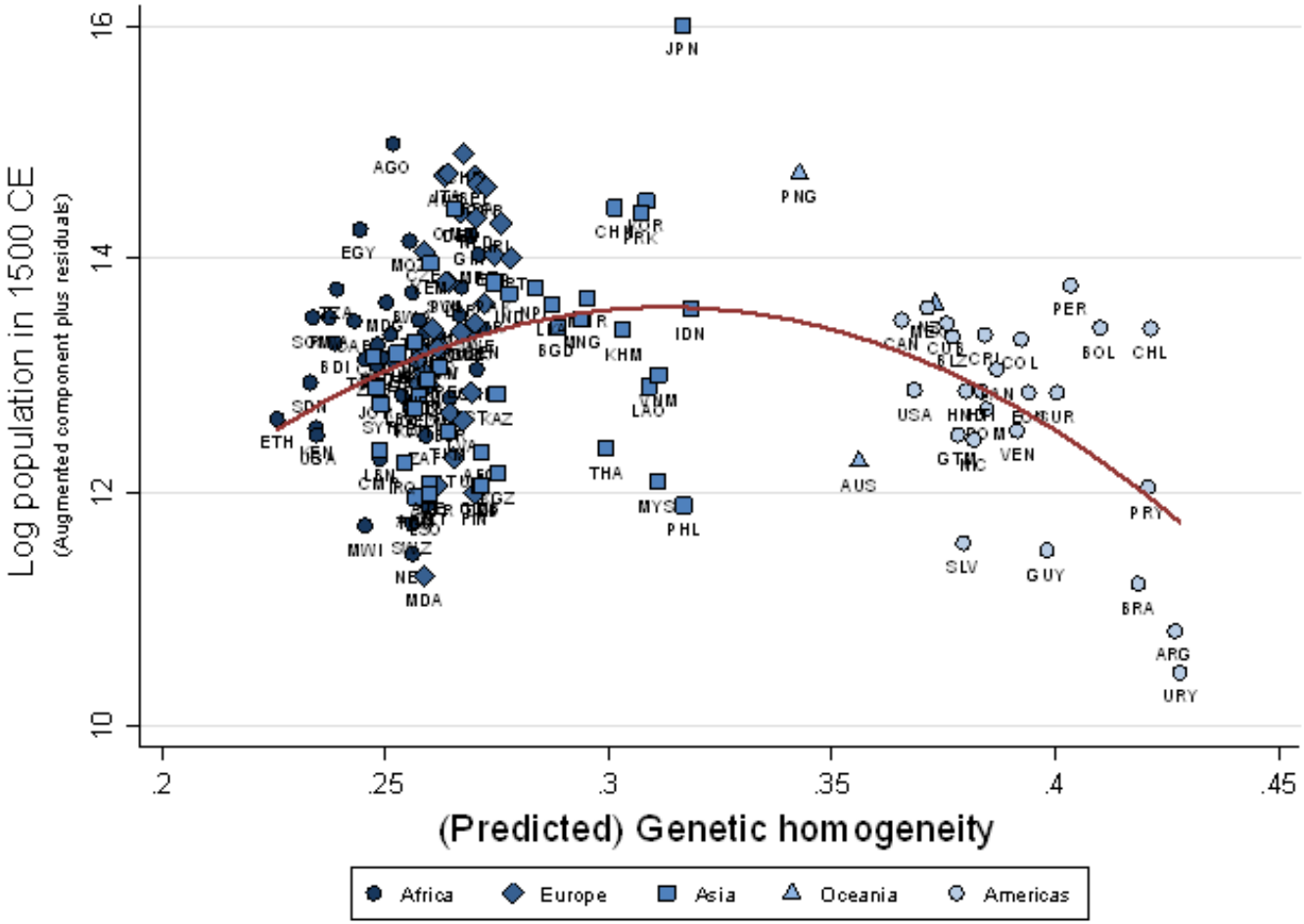


Diversity and Population in 1500 CE

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	OLS	OLS	OLS	OLS
	Dependent Variable is Log Population in 1500 CE					
Predicted diversity	195.327**		165.442**	195.911***	187.905***	175.870**
	(78.400)		(73.372)	(60.893)	(56.122)	(79.292)
Predicted diversity Square	-139.327**		-119.143**	-139.028***	-135.047***	-135.062**
	(57.375)		(53.851)	(44.650)	(40.996)	(55.925)
Log Neolithic transition timing		0.980***	0.836***		1.017***	1.296***
		(0.228)	(0.229)		(0.140)	(0.241)
Log arable land				0.729***	0.735***	0.722***
				(0.067)	(0.050)	(0.050)
Log absolute latitude				-0.259***	-0.451***	-0.447***
				(0.084)	(0.086)	(0.106)
Log land suitability for agriculture				-0.055	-0.027	-0.055
				(0.076)	(0.062)	(0.067)
Continent fixed effects	No	No	No	No	No	Yes
Observations	145	145	145	145	145	145
R^2	0.08	0.12	0.16	0.61	0.72	0.74

Notes: Bootstrapped standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Diversity and Population in 1500 CE



Institutional, Cultural, and Human Capital Variables

- **Social Infrastructure:** An index that quantifies the wedge between private and social returns to productive activities, computed as the average of two separate indices
  - Government anti-diversion policy (GADP) index, (based *International Country Risk Guide)*: the average across five categories, each measured as the mean over the period 1986–1995: (i) law and order, (ii) bureaucratic quality, (iii) corruption, (iv) risk of expropriation, and (v) government repudiation of contracts.
  - An index of openness. The fraction of years in the period 1950–1994 that the economy was open to trade: (i) nontariff barriers cover less than 40% of trade, (ii) average tariff rates are less than 40%, (iii) any black market premium was less than 20% during the 1970s and 80s, (iv) the country is not socialist, and (v) the government does not monopolize over major exports.

Institutional, Cultural, and Human Capital Variables

- **Democracy:** The 1960–2000 mean of an index that quantifies the extent of institutionalized democracy, as reported in the *Polity IV* data set. The *Polity IV* democracy index for a given year is an 11-point categorical variable (from 0 to 10) that is additively derived from *Polity IV* codings on the (i) competitiveness of political participation, (ii) openness of executive recruitment, (iii) competitiveness of executive recruitment, and (iv) constraints on the chief executive.
- **Executive Constraints:** The 1960–2000 mean of an index, reported annually as a 7-point categorical variable (from 1 to 7) by the *Polity IV* data set, quantifying the extent of institutionalized constraints on the decision-making power of chief executives.

Institutional, Cultural, and Human Capital Variables

- **Legal Origins:** A set of dummy variables that identifies the legal origin of the Company Law or Commercial Code of a country. The five legal origin possibilities are: (i) English Common Law, (ii) French Commercial Code, (iii) German Commercial Code, (iv) Scandinavian Commercial Code, and (v) Socialist or Communist Laws.
- **Major Religion Shares:** A set of variables that identifies the percentage of a country's population belonging to the three most widely spread religions of the world. The religions identified are: (i) Roman Catholic, (ii) Protestant, and (iii) Muslim.