1-1-2013

Does Human Genetic Diversity Affect Net Productivity?

Breyon Jontae Williams
University of South Carolina

DOES HUMAN GENETIC DIVERSITY AFFECT NET PRODUCTIVITY?

by

Breyon J. Williams

Bachelor of Arts
University of South Carolina, 2012

Submitted in Partial Fulfillment of the Requirements

For the Degree of Master of Arts in

Economics

Darla Moore School of Business

University of South Carolina

2013

Accepted by:

John H. McDermott, Director of Thesis

Janice B. Breuer, Reader

Lacy Ford, Vice Provost and Dean of Graduate Studies

ABSTRACT

Quamrul Ashraf and Oded Galors' (2013) study, "The 'Out of Africa' Hypothesis, Human Genetic Diversity, and Comparative Economic Development", seeks to explain cross-country variations in economic development, particularly per capita income, through variations in human genetic diversity. Their analysis depends on two fundamental assumptions; genetic diversity's positive effect upon technological productivity and its negative effect upon social capital. This study tests the validity of the results presented by Ashraf and Galor. Specifically, this study seeks to test whether or not the hump-shape relationship observed between income per capita and predicted genetic diversity is validated. Our empirical work supports their findings.

TABLE OF CONTENTS

List of Tables

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

This analysis seeks to test the validity of the "Out of Africa" hypothesis presented by Quamrul Ashraf and Oded Galor. The primary notion of their study is that cross-country variations in economic development can be explained, to a certain degree, by the variation in human genetic diversity (referred to as genetic diversity from this point forward) among national populations. Genetic diversity affects the economic development potential of a country through two counteractive forces: innovation and social capital. First, increases in diversity lead to knowledge creation which positively affects the technological production process. Second, as diversity levels continue to increase, non-cooperation resulting from deterioration in social capital leads to negative effects on the production process. Ashraf and Galor test their hypothesis empirically through OLS estimation using the log of per capita income as their dependent variable. If productivity and genetic diversity behave the way in which they predict, we should observe a hump-shape relationship between per capita income and genetic diversity because of the counteractive forces; their results indicate we do. The mission of this paper is to delve further into the logic of their analysis in an effort to test its validity. The logic this paper seeks to test is whether or not counteractive forces on net productivity are observed with rising levels of genetic diversity in a given population.

In order to test the validity of the Ashraf and Galor case, specifications are constructed in an effort to provide an expanded empirical analysis of the underlying premise in their hypothesis. First, we will test the relationship between genetic diversity and *net productivity.* We define "net productivity" as the product or interaction of the positive technological effect and the negative social capital effect on productivity. Net productivity differs from per capita income in that it ignores the contribution associated with factors of production. Ashraf and Galor claim that higher levels of genetic diversity likely lead to knowledge creation and thus technological innovations in the production process. This creates a positive relation between genetic diversity and production but at diminishing rate. But, genetic diversity exerts a counteractive (negative) effect on productivity because of its effect on social capital. Key to understanding the hump-shape relationship is the idea that genetic diversity has a positive but diminishing marginal effect on productivity which may be offset by genetic diversity's increasingly negative effect on social capital and productivity.

Second, we will directly test the relationship between social capital and genetic diversity. Ashraf and Galor explain the negative force, which becomes prevalent at higher levels of genetic diversity, through the erosion of a population's trust between one another. The rationale governing the effect is that individuals are less likely to trust institutions and fellow citizens if they differ enough to some degree. If this is true, we should observe a significant negative relationship between social capital and genetic diversity. The relationship testing social capital and genetic diversity is tested to reveal the underlying effects present upon net productivity as a result of genetic diversity.

The paper is organized as follows: Chapter 2 will provide further detail regarding the study conducted by Ashraf and Galor including describing the fundamental model constructed for productivity and output per worker as well as describing the approach taken within this study in terms of understanding genetic diversity. Chapter 2 also provides a literature review of relevant studies concerning factors associated with economic variations across countries, particularly those of a prehistoric nature.

Chapter 3 will describe in detail the empirical model employed in this study. A detailed methodology of how our dependent variable, net productivity, *B*, is provided. The empirical study will test the relationship between net productivity and genetic diversity using both the constructed net productivity measurements as well as those measurements used in Hall and Jones (1998) in an attempt to solidify the findings. A justification of Ashraf and Galors' estimate of genetic diversity is also presented within this chapter. In addition to the previous specification regressing net productivity on genetic diversity, cultural variables on respect for others and responsibility using data from Breuer and McDermott (2012) are added. The inclusion of these variables offers an expansion to the single proxy variable for 'trust' employed in Ashraf and Galors' study. The combined effect of respect and responsibility exhibit a significant impact on the aggregate production process relative to trust. Chapter 3 also provides an analysis of cross-country net productivity measurements relative to the U.S.; rankings further feed into the importance of understanding determinants of economic growth.

Chapter 4 will provide an overview of the results from our empirical study. A theoretical critique of Ashraf and Galors' study is presented and describes implications, if any, for their findings. The section also provides suggestions for improving the current study as well as a brief recommendation regarding important variables to utilize when studying comparative economic development. Chapter 5 is the conclusion of the paper.

CHAPTER 2

THE ‘OUT OF AFRICA’ HYPOTHESIS ANALYSIS

2.1 Benchmark Literature Review

Quamrul Ashraf and Oded Galor, authors of “The ‘Out of Africa’ Hypothesis, Human Genetic Diversity, and Comparative Economic Development” (2013), argue that “deep-rooted factors, determined tens of thousands of years ago, have had a significant effect on the course of economic development from the dawn of humankind to the contemporary era”. Ashraf and Galor acknowledge that prevailing hypotheses concerning comparative economic development among countries have traditionally centered around geographical, institutional, and cultural factors, human capital, globalization, colonialism, and ethno-linguistic fractionalization; inclusion of prehistoric determinants are, as argued by Ashraf and Galor, relevant in terms of explaining “the remarkable inequality in income per capita across the globe”. Genetic diversity serves as their prehistoric determinant.

Prior to revealing the results of their study, it is important to describe the fundamental model governing their hypothesis. First, a basic model is constructed detailing the level of gross productivity, A, as a function of institutional, geographical, and human capital factors, Z, alongside the level of diversity, G. This model serves as the basis for the current empirical study in testing whether or not genetic diversity serves as an important determinant for explaining variations in net productivity levels across countries. The model is as follows:

$$A = A\,(\mathrm{Z}, \mathrm{G}) \qquad [1]$$

where $A = A(\mathrm{Z,G}) > 0$, $\frac{\partial A(\mathrm{Z,G})}{\partial \mathrm{G}} > 0$, and $\frac{\partial^2 A(\mathrm{Z,G})}{\partial \mathrm{G}} < 0$ for all A ∀ (0,1). Equation [1] is described as being a positive measurement which experiences decreasing positive marginal effects as a result of increases in diversity.

Assume, now, that a share, δG, of any given country's potential gross productivity is foregone as a result of a lack of cooperation and the resulting inefficiencies in the aggregate production process. Output per worker is then determined by the amount of production factors employed, K and H, the level of gross productivity, *A*, and the level of inefficiency within the production process, δ ∀ (0, 1). The resulting model is as follows:

$$y = (1\text{-}\delta\mathrm{G})A(\mathrm{Z,G})f(\mathrm{K,H}) \qquad [2]$$

where $\frac{\partial y}{\partial \mathrm{G}} > 0$, and $\frac{\partial^2 y}{\partial \mathrm{G}} < 0$. The diminishing marginal effect of diversity on output per worker is justified by Ashraf and Galor (2013) on the grounds of counteractive forces present as diversity within a given population increases. First, a larger array of genetic traits increases the probability of knowledge creation leading to advancements in technological processes for production; an economy's production possibilities frontier is thus expanded as a result. Second, genetic heterogeneity increases the probability that the prevalence of mistrust will increase as populations become more diverse; lower productivity is thus associated with higher degrees of diversity. Based on equation [2], a hump-shape relationship between output per worker and genetic diversity can be observed if net productivity, *B*, also has the same hump-shape relationship with genetic diversity. Net productivity is defined as follows:

$$B = (1\text{-}\delta G)A(Z,G) \qquad [2.1]$$

Equation [2.1] describes the construction of net productivity, *B*, the dependent variable within our analysis.

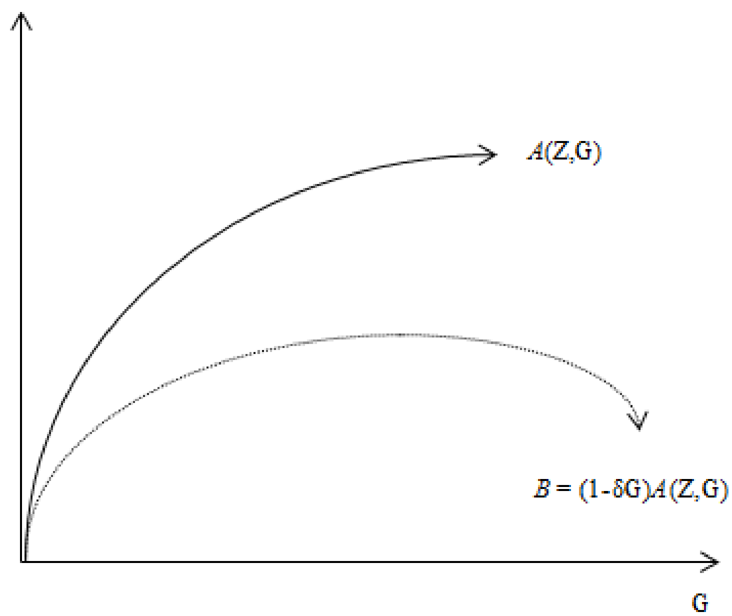

FIGURE 2.1:
Genetic Diversity's Effect on the net productivity resulting from the diminishing marginal return on gross productivity and the negative effect on social capital as discussed in Ashraf and Galor (2013);

The arguments presented above are illustrated in Figure 2.1. The upper curve displays gross productivity as defined in equation [1]. Notice that the marginal effect on gross productivity decreases as genetic diversity increases. The lower curve displays net productivity, *B*. A hump-shape relationship exists because of the counteractive forces observed on net productivity: negative effects on the aggregate production process from the foregone productivity, (1-δG) alongside the positive effects on productivity resulting from technological innovation as diversity levels increase, *A*(Z,G).

The results from Ashraf and Galor show that the degree of diversity within a given population has a hump-shaped effect on economic developmental outcomes; an optimal degree of diversity exists. Figure 2.3, from Ashraf and Galor, displays the hump-shape relationship observed.

Ashraf and Galors' study depends on the correlation between migratory distance from East Africa, described as the "cradle of humankind", and the degree of genetic diversity. The authors state that "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 of their parental colonies". The association between migratory distance and genetic diversity is used to justify and construct the predicted diversity measurements used within their contemporary analysis. Figure 2.2 displays the correlation.

Regarding the empirical model, a cross-sectional analysis is employed using OLS estimation to examine the impact of genetic diversity on log income per capita in the year 2000 CE. Explanatory controlled variables include: Neolithic transition timing and land productivity channels. Further controlled institutional, cultural, and geographical variables are employed to test the robustness of the main specification. Their model is as follows:

$$ln\, y_i = \alpha_0 + \beta_1 \hat{G}_i + \beta_2 \widehat{G_i^2} + \beta_3\, ln\, T_i + \dddot{\beta}_4\, ln\, X_i + \dddot{\beta}_5 \Lambda_i + \dddot{\beta}_6\, ln\, F_i + v_i \qquad [3]$$

where y_i represents the per capita income in a given country i in the year 2000 CE, $\widehat{G}_i$ is the predicted index of contemporary genetic diversity for country i, T_i and X_i are the Neolithic timing and land productivity controls for country i, Λ_i is the vector for institutional and cultural controls for country i, F_i serves as the vector of other geographical controlled variables, and v_i is the error term representing unobserved factors within country i (these variables are discussed further in Chapter 4). Since it was not possible to observe genetic diversity in all countries, the authors predicted the diversity measurements for those unobserved countries through the use of a prior regression of observed diversity measurements on migratory distance from East Africa.

The results of Ashraf and Galors' study indicate a highly significant hump-shape effect of genetic diversity on per capita income within the contemporary period. That is, $\beta_1 > 0$ and $\beta_2 < 0$ in equation [3]. The findings reveal that the optimal level of diversity with respect to the contemporary analysis is higher relative to the optimal level obtained in the historical analysis using population density as a proxy variable for economic development based on the Malthusian perspective. Although the empirical results are arguably quite interesting, the study conducted by Ashraf and Galor has faced its share of criticism.

2.2 Academic Response to Ashraf and Galor

Guedes et al., in the article "Is Poverty in Our Genes?", argue against the claim of Ashraf and Galors' empirical study that "the high degree of diversity among African populations and the low degree of diversity among Native American populations have been a detrimental force in the development of these regions". Their critique is based on the belief that the argument presented by Ashraf and Galor is flawed factually as well as methodologically. The authors state that the study is flawed in three primary ways: First, the consistent misuse of scientific terminology and concepts, particularly as it relates to the relationship between migratory distance and genetic diversity, Second, factual errors in the data, and Third, the inconsistency between their theory and the findings in anthropology, genetics, and sociology on human evolution, cooperation, and innovation.

Regarding the authors' first concern involving the misunderstanding of scientific terminology and concepts, they argue that the use of 'migratory distance' as a proxy variable for the sequential series of founder effects is only appropriate on a continental scale. The authors state that asserting the claim that 'migratory distance' to various settlements across the globe affects genetic diversity is misleading as this was instead influenced by the sequential series of founder effects; geographic distance serves as a proxy for these founder effects. As a result of their critique, the authors suggest that Ashraf and Galor are working with only four data points: Africa, Europe, Asia, and the Americas; the 'predicted genetic homogeneity' estimates for sub-continental populations have no demonstrated scientific basis. Continuing with the second concern, the authors claim that some of the controlled variables employed by Ashraf and Galor such as prehistoric population densities and geographic control factors are poorly chosen. As an example, the author's mention the fact that Ashraf and Galor derive their population estimates from McEvedy and Jones (1978); "a poor and outdated source".

The authors believe there to be justifiable reasons for deeming any data used by Ashraf and Galor for population estimates in the Americas prior to 2000 CE unconnected with reality. Additionally, the authors take up issue with other controlled variables such as the ‘Neolithic transition timing’ variable as well as the variable measuring land suitability for agriculture arguing that, with respect to the former, Ashraf and Galor use data “from Putterman (2008), a source that does not take into account current data and debates in the field”. With respect to the latter, the authors argue that problems exist with how Ashraf and Galor “have ‘corrected’ for land suitability for agriculture”. Lastly, the authors criticize Ashraf and Galor for what they deem “simplistic assumptions about the nature of human behavior”. Recall in the overview of Ashraf and Galors’ study, the authors hypothesize that there exist counteractive forces as diversity levels increase resulting in the diminishing marginal effect of diversity on cooperation and, therefore, production. Guedes et al. criticize Ashraf and Galors’ analysis of human behavior with respect to cooperation and innovation. The authors argue that, based on recent analyses, “evidence indicates that close genetic relationships are not requisite for sustained cooperation among humans”. Turning their attention to the issue of innovation, the authors state that “using the number of scientific articles published per year, per capita” as a way to test for a relationship with the predicted genetic diversity values could be problematic. As stated by Guedes et al., “the number of scientific articles published by a nation is closely tied to a nation’s history… additional factors likely include the amount of government funding allocated to research and high degrees of economic specialization”; the variables employed such as ‘years of schooling’ are inadequate controls for the underlying factors which explain variations in innovation (2013).

2.3 Related Literature

Rosenberg (2002) and colleagues were the first to utilize the Human Genome Diversity Project (HGDP) collection, emphasizing the importance of geographical isolation in determining genetic divergence (Cavalli-Sforza 2005). The study conducted by Ashraf and Galor extends the analysis; exploiting the relationship between migratory distance and genetic diversity in order to construct predicted values for genetic diversity to serve as estimates for unobserved country data in the contemporary model. Recall the critiques presented in the above section: the use of 'migratory distance' as a proxy variable for the serial founder effects, poorly chosen controlled variables, and the role of genetic diversity as it relates to cooperation and innovation; there exists some related literature which may serve to justify some of the assumptions presented by Ashraf and Galor. Of particular importance is the use of migratory distance in explaining cross country variation of genetic differentiation.

Ramachandran et al. in a study concerning the relationship of genetic diversity and geographical distance in human populations reveal a correlation of geographic distance and genetic differentiation (as measured by F_{st}). Figure 3.1 displays the correlation exploited by Ashraf and Galor. In addition to the correlation observed, their study finds that expected heterozygosity among populations, from a global representative data set, are best explained by an expansion originating in Africa; no other geographic origin external to Africa accounts as well for the observed patterns of genetic diversity.[1] In response to the critiques of Guedes et al. regarding Ashraf and Galors' use of 'migratory distance' to estimate genetic diversity, the study by Ramachandran et al. states

[1] Ashraf and Galor (2013) define expected heterozygosity as the probability that two individuals, selected at random from the relevant population, are genetically different from one another.

that “although the relationship between F_{st} and geographic distance has been interpreted in the past as the result of an equilibrium model of drift and dispersal, simulation shows that the geographic pattern of heterozygosities… is consistent with a model of serial-founder scenario, the relationship between genetic diversity and geographic distance allows us to derive bounds for the effects of drift and natural selection on human genetic variation. Ramachandran et al. further extend the interpretation of their results stating that “an expansion of modern humans outward from a single center is an alternative way of producing a global correlation between geographic and genetic distances. Geographical expansion events may have happened in many small steps, with each such migration involving a sampling from the previous subset of the original population. This sampling would have led to a stepwise increase in genetic drift and a concomitant decrease in genetic diversity; a serial founder effect”.

Overall, the argument presented by Ramachandran et al. justifies the reasoning behind Ashraf and Galors’ use of ‘migratory distance’ to estimate genetic diversity measurements within a given population.

Regarding related literature, Ashraf and Galor, in their article “Isolation and Development”, continue in their efforts towards employing prehistoric factors within empirical models in order to explain the course of comparative economic development. Ashraf and Galor argue that prehistoric geographical isolation among countries resulted in a consistent positive effect on the process of development contributing to the contemporary differentiation in the cross-country standard of living. In justifying their argument, Ashraf and Galor state that “the diminished ability of geographically isolated societies to benefit from advancements in the world technological frontier may have induced an independent process of technological advancements, fostering a long-lasting cultural environment conducive to innovations" (2013). The 'innovation' assumption is necessary if the hypothesis is to be justified theoretically. Similar to the "Out of Africa" hypothesis, Ashraf and Galor construct both an historical and contemporary analysis; dependent variables for the two analyses are population density in 1900 and per capita income in 2000 respectively.

The results of their empirical study reveal a significant positive coefficient estimate on the isolation variable (measured as an index reflecting the average time required to travel from the capital of a country to each square kilometer of land on the surface of the earth, incorporating land routes that minimize travel time in the absence of maritime and airborne transportation technologies). The empirical analysis presented here could support further efforts to adjust the conventional methodology utilized when determining influential factors affecting the economic development of a country.

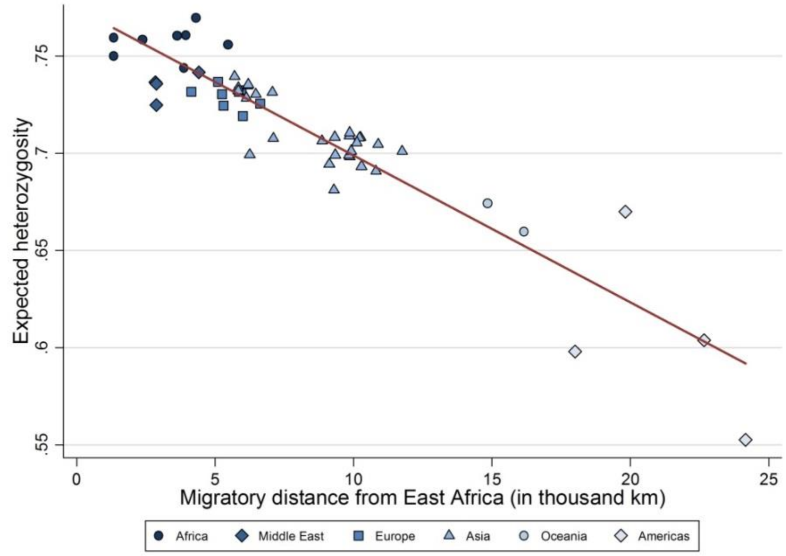

FIGURE 2.2
Expected Heterozygosity and Migratory Distance from East Africa; depicts the inverse relationship between genetic diversity and distance from East Africa

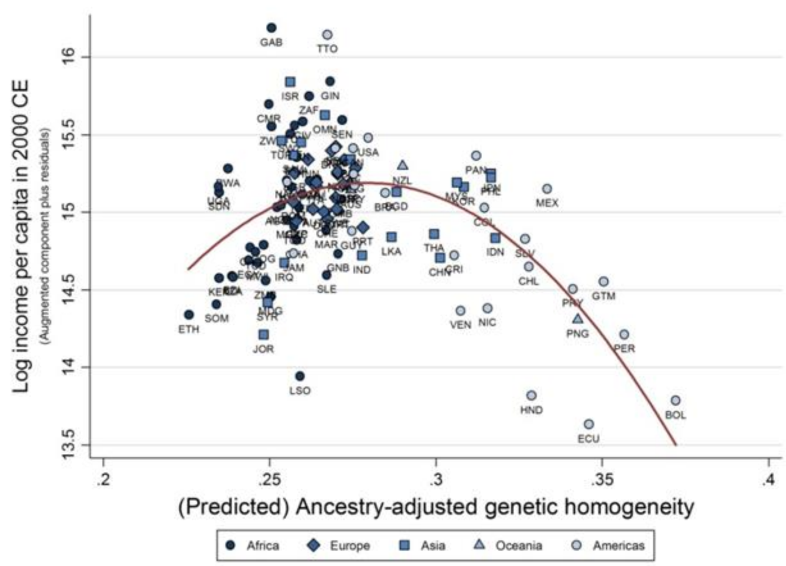

FIGURE 2.3
Ancestry-Adjusted Genetic Homogeneity and Income per Capita in 2000 CE;
depicts the hump-shape relationship of genetic diversity upon log per capita income in 2000 CE resulting from the counteractive forces discussed in Chapter 2. See equation 2. Source: Ashraf and Galor: "The 'Out of Africa Hypothesis', Human Genetic Diversity, and Comparative Economic Development."

2.4 Justifying Genetic Diversity's Effect on the Production Process: A Theoretical Explanation

Ashraf and Galor seek to explain the amount of variation in cross-country economic development through the variation in human genetic diversity among national populations. The hump-shape relationship observed between log income per capita in 2000 CE and genetic diversity is further justified based on their theoretical assumption regarding the relationship between human behavior and genetic diversity (see Figure 2.3). Based on their hypothesis, "heterogeneity raises the likelihood of disarray and mistrust, reducing cooperation and disrupting the socioeconomic order. Higher diversity is therefore associated with lower productivity…" the authors proceed in their explanation indicating that "a wide spectrum of traits is more likely to contain those that are complementary to the advancement and successful implementation of superior technological paradigms. Higher diversity therefore enhances society's capability to integrate advanced and more efficient production methods…" (2013); the counteractive forces present as genetic diversity becomes more prevalent within a given population results in the decreasing marginal effect observed on net productivity. Although the empirical results presented by Ashraf and Galor support their hypothesis, is the theoretical explanation reflective of reality? This question warrants attention.

First, let's focus on the negative effect on net productivity resulting from higher levels of diversity. Matt McGue and Thomas J. Bouchard, Jr, authors of "Genetic and Environmental Influences on Human Behavioral Differences", place in perspective the relationship between one's genetic make-up and their behavior. Based on their study, the genetic make-up of an individual is likely to influence an array of observable characteristics related to human behavior. Characteristics include cognitive abilities as well as personality and interests. As stated by McGue and Bouchard, "the most widely utilized scheme for characterizing personality traits is the Big Five – extraversion, agreeableness, conscientiousness, neuroticism, and openness" (1998). Although much controversy remains within the field of behavioral genetics, there exists a mutual understanding that genetic differentiation can explain, to some degree, the variation observed in human behavior. This particular article is important with respect to the current analysis because it provides some realistic probability that genetic diversity can lead to variations in the capabilities of individuals within a population. It is not the purpose of this paper to provide a detailed overview regarding existing debates within the field of human behavioral genetics; only to present a connection with genetic differentiation and observable differences among individuals is the goal of this analysis. Thus, if individuals in a given population are able to observe differences between themselves and others then the idea of "disarray and mistrust" could be a likely event. Second, if genetic diversity leads to variations in human behavior, particularly with regards to cognitive abilities, then it is also plausible that complimentary collaborations could become more prevalent as diversity levels increase leading to technological advancements in the aggregate production process.

Previous studies have emphasized the effect of identity variation between and within groups on economic growth; particularly ethnic and religious fractionalization. Montalvo and Reynal-Querol (2004) studied the relationship between ethnic diversity and economic development. Their analysis compared different measurements of ethnic diversity; ethnic fractionalization compared with ethnic polarization.

Their study confirmed that ethnolinguistic fractionalization is inversely related with growth but not through indirect channels such as investment, public consumption and the incidence of civil wars; widely cited as reasons for the negative relationship. Montalvo and Reynal-Querol provide an important theoretical explanation for how differences within populations can result in negative growth. First, cleavages result in frictions among sub-groups in a population. As stated by the authors, “when the society is divided by religious, ethnolinguistic, or racial differences, tensions emerge along these divisions.” This is important to note because genetic differentiation among populations increases the likelihood that differences such as those mentioned by Montalvo and Reynal-Querol are observed. The authors proceed in their explanation stating that “resources spent by the groups in order to obtain political influence can be considered as a social cost with a negative effect on economic growth because it implies a nonproductive use of these inputs… the government will increase its consumption in order to mitigate potential conflict, which also has a negative effect on growth”. These explanations are likely events hypothesized by Ashraf and Galor as genetic diversity increases to higher levels; theoretically their hypothesis seems plausible.

CHAPTER 3

THE EMPIRICAL ANALYSIS: DATA AND STRATEGY

3.1 The Empirical Model

Focusing specifically on genetic diversity's effect on net productivity, the regression specification will mirror that of equation [3] employed by Ashraf and Galor. The main difference between our study and Ashraf and Galor is that we use net productivity, *B*, as the dependent variable, instead of income per capita. The empirical study performed by Ashraf and Galor tested indirectly the effect of genetic diversity on productivity through the byproduct effect on income per capita. Based on equation [2.1], the current results should reveal a significant hump-shape relationship between net productivity and genetic diversity assuming their hypothesis is correct; diminishing marginal effect should be observed. In an effort to test the validity of their results, it is important that the current specification preserve much of their controlled variables so as to ensure that the results are not a byproduct of a completely different model. Thus the model is as follows,

$$ln\, B_i = \alpha_0 + \beta_1 \hat{G}_i + \beta_2 \widehat{G_i^2} + \beta_3 \ln T_i + \dddot{\beta}_4 \ln X_i + \dddot{\beta}_5 \Lambda_i + \dddot{\beta}_6 \ln F_i + \varepsilon_i \qquad [4]$$

where B_i, the net productivity measurement of country *i* in the year 2000 CE, equals (1-δG)*A*. The remaining variables are relatively unchanged from equation [3].

In addition to the main specification, this study will take into account the critiques presented in Chapter 2 by Guedes et al. regarding genetic diversity's effect on innovation and cooperation; an additional specification will be employed to retest the relationship between genetic diversity and cooperation (see Ashraf and Galors' The Cost and Benefits of Genetic Diversity: The "Out of Africa" Hypothesis: page 39).

3.2 The Genetic Diversity Measurement

Predicted genetic diversity serves as the focus explanatory variable in the current specification. Data for genetic diversity is obtained from Ashraf and Galor (2013). A primary criticism of Guedes et al. regarding Ashraf and Galors' study was the measure of genetic diversity at the country level. First, we describe the methodology behind measuring genetic diversity and second, justify the use of this measurement in the current model. The Ashraf and Galor study constructs what is known as the index of genetic diversity for contemporary national populations. The authors employ the concept of F_{st} genetic distance in order to estimate genetic diversity across countries. Assume in country A there exists two different groups of people, those who are silver (S) and those who are gold (G). The F_{st} genetic distance between the silver and gold groups measures the ratio of their combined genetic diversity that is not explained by the weighted average of their respective genetic differentiation. Thus F_{st} is calculated as,

$$F_{st}^{SG} = 1 - \left(\frac{\delta_S H_{exp}^{S} + \delta_G H_{exp}^{G}}{H_{exp}^{SG}}\right) \qquad [5]$$

where H_{exp} denotes the genetic diversity of the respective ethnic group, δ denotes the share in the overall population of the respective ethnic group, and H_{exp}^{SG} is the expected heterozygosity of the respective population; *SG* in this case. The generalized formula measuring the F_{st} for N sub-populations within country A follows as,

$$F_{st}^{X_n} = 1 - \left(\frac{\sum_{i=0}^{n} \delta_{X_i} H_{exp}^{X_i}}{H_{exp}^{X_n}} \right) \qquad [5.1]$$

where X_i denotes group *i* within the population. Based on the Ashraf and Galor study, $H_{exp}^{X_n}$ is the variable that is estimated; human genetic diversity of a given population. Solving for $H_{exp}^{X_n}$ results in the following equation:

$$G = H_{exp}^{X_n} = \left(\frac{\sum_{i=0}^{n} \delta_{X_i} H_{exp}^{X_i}}{1 - F_{st}^{X_n}} \right) \qquad [5.2]$$

Ashraf and Galor indicate that the calculations for cross-country population diversity would be limited because the HGDP-CEPH provides sample data for only 53 ethnic groups. In addition to the limitation on expected heterozygosity among ethnic groups, their study also lacks data on genetic distances between groups, F_{st}. The genetic diversity measurement employed by Ashraf and Galor thus exploit the predictive power of migratory distance from East Africa to estimate the expected heterozygosity at the ethnic group level while appealing to the serial founders effect to estimate F_{st}. Based on data from Ramachandran et al. (2005), there exists a strong positive effect of pairwise migratory distance on pairwise genetic distance across all pairs of ethnic groups within the HGDP-CEPH sample (Ashraf and Galor 2013). Figure 3.1 illustrates the strong positive correlation between pairwise migratory distance and pairwise genetic distance; this relationship is exploited in order to expand the available F_{st} measurements in estimating the unobserved genetic diversity levels. The methodology employed by the authors does provide support for the use of their variable as an appropriate measurement for genetic diversity.

In order to estimate $H_{exp}^{X_n}$, or G, $F_{st}^{X_n}$ must first be estimated to ensure a measurement exists for those countries which are not observed. Thus, the authors run a regression of the following relationship:

$$F_{st}^{X_n} = \alpha + \beta_1 P^n + other\ factors \qquad [5.3]$$

Equation [5.1] is estimated to obtain the coefficient estimates, $\hat{\alpha}$ and $\widehat{\beta_1}$, in order to utilize $\widehat{F_{st}^{X_n}}$ in equation [5.2]. Once the pairwise genetic distance measurement is estimated, Ashraf and Galor exploit the correlation between expected heterozygosity and migratory distance from East Africa, illustrated in Figure 2.1. Thus, the authors run a regression of the following relationship:

$$H_{exp}^{X_i} = \alpha + \beta_1 D_i + other\ factors \qquad [5.4]$$

Equation [5.4] is estimated to obtain the coefficient estimates, $\hat{\alpha}$ and $\widehat{\beta_1}$, in order to utilize $\widehat{H_{exp}^{X_i}}$ in equation [5.2]. Once these two measurements, $\widehat{F_{st}^{X_n}}$ and $\widehat{H_{exp}^{X_i}}$, are obtained, we can now solve for the genetic diversity of any given country whether or not the measurements were initially observed. Thus, the genetic diversity measurement observed in equations [3] and [4] is predicted as follows:

$$\hat{G} = \widehat{H_{exp}^{X_n}} = \left(\frac{\sum_{i=0}^{n} \delta_{X_i} \widehat{H_{exp}^{X_i}}}{1 - \widehat{F_{st}^{X_n}}} \right) \qquad [5.5]$$

Currently, there are few alternative measures for genetic diversity. Ashraf and Galors' measurement serves as a viable option to include in the current specification. Also, the primary purpose of this paper is to test the validity of the results presented in their analysis with respect to net productivity, so utilization of an alternative measurement for genetic diversity would be non-informative.

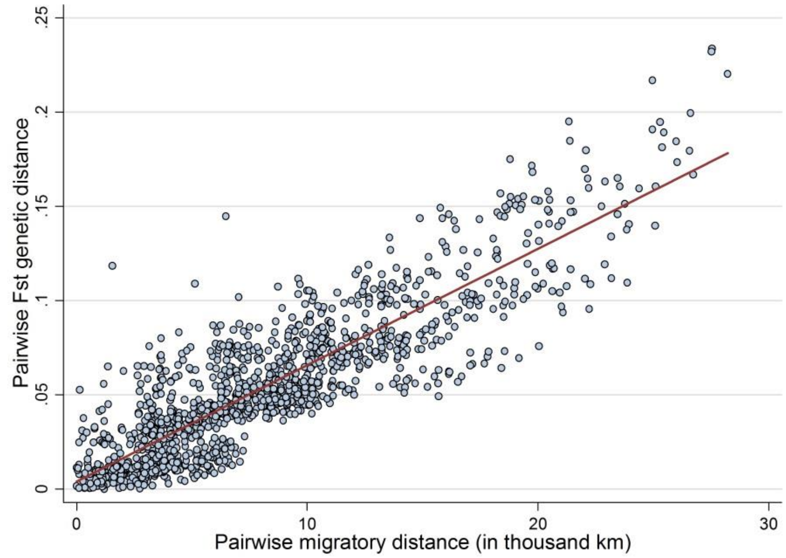

FIGURE 3.1:
Displays the relationship between the Pairwise F_{st} genetic distance and Pairwise migratory distance; the serial founder's effect reveals the theoretical explanation justifying the measurement of genetic diversity employed by Ashraf and Galor (2013). Source: Ashraf and Galor: "The 'Out of Africa Hypothesis', Human Genetic Diversity, and Comparative Economic Development."

3.3 The Net Productivity Measurement

In order to employ the specification described in equation [4], a variable measuring the cross-country net productivity level must be constructed. Hall and Jones, in the study "Why Do Some Countries Produce So Much More Output per Worker than Others?", calculated the level of productivity directly from a Cobb-Douglas production function using data on output, capital, and educational attainment across countries in 1988. The net productivity measurement employed in the current specification will be constructed based on their methodology for the year 2000 CE. The production function, based on Hall and Jones, is written in terms of output per worker as,

$$y_i = B_i \left(\frac{K_i}{Y_i}\right)^{\frac{\alpha}{1-\alpha}} h_i \qquad [6]$$

where $y_i = \left(\frac{Y_i}{L_i}\right)$, $h_i = \left(\frac{H_i}{L_i}\right)$ and $H_i = e^{\varphi E_i} L_i$. Equation [6] decomposes output per worker into the capital-output ratio, educational attainment and net productivity; where K_i is the physical capital stock, Y_i is aggregate production, H_i is the amount of human capital-augmented labor used in the production process, L_i is the number of workers, and B_i is the labor-augmented measure of net productivity, the measurement employed in equation [4] for country *i*.

In calculating the measure of net productivity, data is collected for all other variables and the net productivity residual is then used as our statistic.

$$B_i = \frac{y_i}{\left(\frac{K_i}{Y_i}\right)^{\frac{\alpha}{1-\alpha}} h_i} \qquad [6.1]$$

To calculate the measure of net productivity, we use data on output, labor input, average educational attainment, and physical capital for the year 2000. Data on output is collected from the World Bank, labor input from the Conference Board, the stock of physical capital from Berlemann and Wesselhöft (2012) and educational attainment from Barro and Lee (2001). The current productivity calculation will follow the neoclassical approach presented in Hall and Jones (1998) where $\alpha = \frac{1}{3}$. The net productivity measurement, B_i, is expressed as a ratio to U.S. values for each country observed.

TABLE 3.1: Net Productivity Rankings: Ratios to U.S. Values – 1988 and 2000

2000						1988					
Rank	Country	B	Rank	Country	B	Rank	Country	B	Rank	Country	B
1	Luxembourg	1.221	32	Poland	0.201	1	Italy	1.207	32	Cyprus	0.646
2	U.K.	1.019	33	Brazil	0.189	2	France	1.126	33	New Zealand	0.631
3	Norway	1.012	34	Guatemala	0.179	3	Spain	1.107	34	Uruguay	0.579
4	United States	1.000	35	Malaysia	0.171	4	Luxembourg	1.098	35	Pakistan	0.566
5	Japan	0.978	36	Tunisia	0.158	5	Canada	1.034	36	Morocco	0.527
6	Denmark	0.938	37	Hungary	0.141	6	U.K.	1.011	37	Costa Rica	0.506
7	Switzerland	0.909	38	Algeria	0.127	7	United States	1	38	Turkey	0.503
8	Iceland	0.864	39	Jordan	0.115	8	Jordan	0.998	39	Malaysia	0.45
9	Sweden	0.860	40	Peru	0.110	9	Austria	0.979	40	Peru	0.409
10	Belgium	0.854	41	Morocco	0.090	10	Belgium	0.978	41	Chile	0.403
11	France	0.843	42	Thailand	0.079	11	Netherlands	0.946	42	Ecuador	0.386
12	Ireland	0.839	43	Pakistan	0.064	12	Iceland	0.933	43	Thailand	0.369
13	Finland	0.827	44	Sudan	0.062	13	Mexico	0.926	44	Senegal	0.361
14	Italy	0.812	45	Ecuador	0.062	14	Sweden	0.897	45	Mozambique	0.321
15	Austria	0.791	46	Bolivia	0.061	15	Switzerland	0.883	46	Bolivia	0.305
16	Canada	0.722	47	Romania	0.052	16	Australia	0.856	47	Hungary	0.293
17	Netherlands	0.702	48	Philippines	0.051	17	Bangladesh	0.844	48	Cameroon	0.274
18	Australia	0.581	49	Indonesia	0.044	18	Venezuela	0.839	49	India	0.267
19	Spain	0.527	50	Cameroon	0.033	19	Algeria	0.771	50	Indonesia	0.242
20	Malta	0.502	51	Bangladesh	0.033	20	Brazil	0.758	51	Poland	0.235
21	Cyprus	0.485	52	India	0.032	21	Portugal	0.755	52	Sudan	0.233
22	Greece	0.472	53	Senegal	0.030	22	Guatemala	0.753	53	Uganda	0.224
23	Portugal	0.447	54	Mali	0.027	23	Malta	0.743	54	Philippines	0.223
24	Uruguay	0.445	55	Kenya	0.026	24	Finland	0.728	55	Mali	0.196
25	Argentina	0.387	56	China	0.026	25	Ireland	0.709	56	Romania	0.18
26	New Zealand	0.374	57	Mozambique	0.020	26	Denmark	0.705	57	Kenya	0.165
27	Venezuela	0.305	58	Uganda	0.016	27	Norway	0.699	58	China	0.106
28	Mexico	0.298	59	Zambia	0.013	28	Tunisia	0.683	59	Zambia	0.079
29	Turkey	0.280				29	Greece	0.674			
30	Chile	0.252				30	Japan	0.658			
31	Costa Rica	0.205		Average	0.389	31	Argentina	0.648		Average	0.621

TABLE 3.2: Output per Worker Decomposed, Relative to the U.S. – 2000

COUNTRY	$\frac{Y}{L}$	$\left(\frac{K}{Y}\right)^{\frac{\alpha}{1-\alpha}}$	$\frac{H}{L}$	B
Algeria	0.087	1.288	0.534	0.127
Argentina	0.308	1.076	0.739	0.387
Armenia	0.021	1.483	0.861	0.016
Australia	0.647	1.204	0.925	0.581
Austria	0.719	1.189	0.764	0.791
Bangladesh	0.017	1.133	0.456	0.033
Belgium	0.791	1.134	0.817	0.854
Bolivia	0.040	0.981	0.673	0.061
Brazil	0.114	1.073	0.562	0.189
Bulgaria	0.056	1.316	0.783	0.054
Cameroon	0.021	1.224	0.514	0.033
Canada	0.684	1.082	0.876	0.722
Chile	0.204	1.084	0.749	0.252
China	0.023	1.476	0.618	0.026
Costa Rica	0.146	1.017	0.699	0.205
Cyprus	0.414	1.071	0.798	0.485
Czech Republic	0.170	1.289	0.928	0.142
Denmark	0.811	1.062	0.815	0.938
Ecuador	0.054	1.359	0.640	0.062
Estonia	0.139	1.413	0.917	0.107
Finland	0.743	1.123	0.800	0.827
France	0.725	1.106	0.778	0.843
Germany	0.670	1.140	0.843	0.697
Greece	0.409	1.172	0.740	0.472
Guatemala	0.080	0.961	0.462	0.179
Hungary	0.153	1.222	0.887	0.141
Iceland	0.778	1.157	0.778	0.864
India	0.018	1.224	0.449	0.032
Indonesia	0.026	1.146	0.504	0.044
Ireland	0.805	1.085	0.884	0.839
Italy	0.674	1.121	0.740	0.812
Japan	1.014	1.209	0.858	0.978
Jordan	0.098	1.282	0.666	0.115
Kazakhstan	0.041	1.540	0.810	0.033
Kenya	0.015	1.034	0.555	0.026
Latvia	0.116	1.269	0.784	0.117
Luxembourg	1.075	1.102	0.799	1.221

TABLE 3.2 continued

COUNTRY	$\frac{Y}{L}$	$\left(\frac{K}{Y}\right)^{\frac{\alpha}{1-\alpha}}$	$\frac{H}{L}$	B
Mali	0.011	1.266	0.319	0.027
Malta	0.378	0.991	0.760	0.502
Morocco	0.051	1.285	0.437	0.090
Mozambique	0.007	1.130	0.313	0.020
Netherlands	0.664	1.098	0.862	0.702
New Zealand	0.399	1.139	0.937	0.374
Norway	1.015	1.109	0.904	1.012
Pakistan	0.028	1.012	0.431	0.064
Peru	0.090	1.171	0.693	0.110
Philippines	0.041	1.138	0.710	0.051
Poland	0.165	1.043	0.789	0.201
Portugal	0.326	1.159	0.629	0.447
Romania	0.048	1.145	0.808	0.052
Russian Federation	0.056	1.606	0.894	0.039
Senegal	0.017	1.233	0.453	0.030
Slovenia	0.306	1.145	0.907	0.294
Spain	0.495	1.223	0.767	0.527
Sudan	0.021	0.860	0.385	0.062
Sweden	0.804	1.072	0.872	0.860
Switzerland	0.870	1.175	0.815	0.909
Tajikistan	0.007	1.062	0.812	0.008
Thailand	0.055	1.278	0.546	0.079
Tunisia	0.094	1.152	0.517	0.158
Turkey	0.169	1.088	0.555	0.280
Uganda	0.009	1.120	0.466	0.016
Ukraine	0.021	1.994	0.852	0.012
United Kingdom	0.750	1.004	0.733	1.019
United States	1.000	1.000	1.000	1.000
Uruguay	0.299	0.945	0.710	0.445
Venezuela	0.196	1.161	0.554	0.305
Zambia	0.010	1.410	0.574	0.013

Table 3.1 displays the net productivity measurements constructed as ratios to U.S. values for those countries in which data was available for both the years 1988 and 2000. The 1988 column reflects the productivity calculations constructed by Hall and Jones (1988) while the 2000 column displays those measurements calculated here based on the Hall and Jones methodology. Many characteristics between the two data sets are similar although they do differ to some degree. It is important to note that the table reflects relative positioning between the observed countries and not the absolute positions of the countries. When one compares the top ten most productive countries in 1988 to the 2000 rankings, only four have remained in that group: Luxembourg, the U.K., the United States, and Belgium.

Of the ten least productive countries in 1988, there are 5 which have managed to remain in that group: Uganda, Mali, Kenya, China, and Zambia. The United States has seen its productivity levels rise between 1988 and 2000 reflected not only within the rankings but also implicitly within the average ratio from 1988 to 2000. In 1988, the average productivity ratio was 62.1 percent; on average a country's productivity level would be 62.1 percent of the U.S. level. The average level has since decreased to 38.9 percent; on average a country's productivity level would be 38.9 percent of the U.S level in 2000. The 2000 data presents a larger range of productivity measures which likely hints toward divergence with respect to productivity; not all countries appear to be benefitting from the technological innovations observed since 1988.

Based on the 2000 productivity measurements, Luxembourg ranks the highest with productivity levels which are 22.1 percent higher than the United States. The United Kingdom and Norway rank as number two and three respectively with productivity levels that are 1.9 percent and 1.2 percent higher than the United States. Comparing the data to alternative statistics shows that our measurements align with conventional variables employed to measure productivity. Governmental agencies, such as the U.S. Bureau of Labor Statistics, consistently measure productivity in terms of real GDP per hours worked. Countries such as the United States, Switzerland, Norway and Belgium consistently rank as some of the most productive in the world. Overall our measurements match well with that of Hall and Jones: of the 20 most productive countries in their 1988 data (in relative terms), 14 of those countries also are ranked in the top 20 in year 2000.

Of particular interest are the countries which have seen their relative positions change drastically overtime. According to the 2000 calculations in Table 1, Japan ranks as the fifth most productive country compared to its position in 1988 as the thirtieth most productive country. Although the calculations themselves provide some insight regarding the relative degree of productivity, one must be cautious in how much to interpret from the data. China is also another interesting country to view because of its rapid economic growth since 1985. China, with the second-largest economy in the world, measures poorly in terms of productivity in both years. Could China's growth be explained by high-intensive capital accumulation? If so, based on the Solow growth model, we could indeed see an economic stagnation in the coming years in China. Understanding why certain countries rank where they do drives the current study at hand. The primary specification will be modeled in order to test the relationship, if any, between productivity and genetic diversity.

CHAPTER 4

RESULTS

4.1 The Regression Models

The primary results are generated based on equation [4]. The Ashraf and Galor model is re-tested with the log of net productivity acting as the dependent variable instead of income per capita. The tables that follow first test the relationship between net productivity and predicted genetic diversity using both the constructed productivity measurements and the Hall and Jones measurements. Second, we refine the specification testing net productivity and predicted genetic diversity controlling for other variables. Lastly, we test the relationship between social capital and predicted genetic diversity. The first two relationship tests allow variation in the functional form of the model; a linear and quadratic relationship is tested alongside the utilization of the logarithmic form of net productivity. Each of the regression models utilizes cross-sectional data.

Regarding the definitions and data sources of key variables, predicted genetic diversity is from Ashraf and Galor (2013), the timing of the Neolithic Revolution is from Putterman (2008), institutional and cultural controls include the social infrastructure index of Hall and Jones (1999), ethnic fractionalization index from Alesina et al. (2003), and legal origin dummies alongside the share of the population affiliated with major world religions from the data set of La Porta et al. (1999). The respect and responsibility data from Breuer and McDermott (2012) is also used in place of the Hall and Jones' measure of social infrastructure. The Neolithic transition timing variable reflects the number of years elapsed, as of the year 2000 CE, since the onset of sedentary agriculture. Tables 4.1, 4.2, and 4.7 display the summary statistics for the respective regressions. The summary statistics are primarily given so that one can verify the magnitudes regarding the size of the coefficient estimate of genetic diversity within any given model.

4.1.1 Net Productivity and Predicted Genetic Diversity

Tables 4.3-4.6 focus exclusively upon the relationship between net productivity and predicted genetic diversity. Six models are presented; each table mirrors the specifications employed in Ashraf and Galor (2013) with the exception that the dependent variable is the log of net productivity. Table 4.3 presents the linear functional form of the relationship between net productivity and genetic diversity. Each of the models reports a significant and positive coefficient estimate of the predicted genetic diversity variable; models 1, 4-5 at the 5 percent significance level and models 2, 3, and 6 at the 10 percent significance level. Model 1 reports that a 1 percent increase in predicted genetic diversity from its mean results in an 11 percent increase in net productivity. The magnitudes can be shown to be relatively large in terms of the marginal change in the predicted genetic diversity measure. In fact, a 1 percent increase in predicted genetic diversity from its mean results in an 8, 7, 9, and 8 percent increase in net productivity respectively. The models also explain an immense amount of the variation in net productivity with R-squared values of .73, .85, .87, .91, .91, .92 respectively. Table 4.3 provides empirical results consistent with equation [1]; the positive effects on productivity dominating when the possibility for the negative effect on social capital is removed.

Recall, based on the analysis presented in chapter 2, we should observe a quadratic relationship between predicted genetic diversity and net productivity; Table 4.4 takes the possibility of such a relationship into account.

TABLE 4.1: Summary Statistics for Tables 4.3-4.4 and 4.8-4.9

Variable	Mean	Std. Dev.	Min	Max
Log net productivity	-1.639	1.361	-4.377	0.019
Predicted diversity (ancestry adjusted)	0.720	0.030	0.628	0.765
Predicted diversity squared (ancestry adjusted)	0.519	0.042	0.394	0.586
Log Neolithic transition timing	8.592	0.364	7.244	9.173
Log percentage of arable land	2.550	0.863	0.993	4.129
Log absolute latitude	3.119	1.008	0.000	4.159
Social infrastructure	0.576	0.269	0.156	1.000
Ethnic fractionalization	0.383	0.274	0.012	0.930
Years of schooling	5.470	2.939	0.409	10.862

Ashraf and Galor predict a hump-shape relationship between genetic diversity and net productivity, thus Table 4.4 utilizes the ‘predicted genetic diversity squared’ variable in order to test the existence of a quadratic relationship. Table 4.4 does not display any significant hump-shape relationship between net productivity and predicted genetic diversity. In fact, models 4-6 do not generate the appropriate signs in order for a hump-shape relationship to exist between the two variables. The optimality levels, .76 and .82, in models 1 and 2 respectively are consistent with levels presented by Ashraf and Galor (2013) although these results are not statistically significant. Tables 4.3 and 4.4 do not present results consistent with the Ashraf and Galor model; we take into account the possibility of statistical inefficiencies resulting from the limited sample size; we now rerun the models using the Hall and Jones productivity measurements which are observed for twice as many countries as our constructed productivity measurement.

TABLE 4.2: Predicted diversity means for Tables 4.5 and 4.6

Variable	Model	Mean	Std. Dev.	Min	Max
Predicted diversity (ancestry adjusted)	1	0.724	0.029	0.628	0.766
	2	0.724	0.029	0.628	0.766
	3	0.724	0.029	0.628	0.766
	4	0.724	0.029	0.628	0.766
	5	0.724	0.029	0.628	0.766
	6	0.721	0.030	0.628	0.765

Table 4.5 re-estimates the model displayed in Table 4.3 using productivity measurements from Hall and Jones. The sample size is twice as large as the sample used to estimate the models of Tables 4.3 and 4.4 since Hall and Jones constructs a larger sample of countries for which the productivity measurement is available. All of the models display a positive coefficient estimate of predicted genetic diversity on net productivity although none of the estimates are statistically significant at any conventional level. The sizes of the coefficient estimates are smaller than that of Table 4.3 in terms of the marginal magnitudes. The results in Table 4.5 show that a 1 percent increase in predicted genetic diversity from the mean results in a .8, 2.5, 2.4, 2.4, 2.4, and 2.2 percent increase in net productivity respectively. Because Table 4.5 uses a larger sample size, its results may be considered to be more robust than those in Table 4.3.

Table 4.6 re-takes into account the possibility of a quadratic relationship between net productivity and predicted genetic diversity. Like the comparison between Table 4.3

and 4.5, the results in Table 4.6 take precedence over the results in Table 4.5. Unlike Table 4.5, each of the six models within Table 4.6 generates a hump-shape relationship between net productivity and predicted genetic diversity although the relationship is not statistically significant at any conventional level. Optimality levels mirror that of the results present within Ashraf and Galor; optimal levels are .71, .79, .78, .75, .75, and .74 respectively. The magnitudes in Table 4.6 are respectively large. A 1 percent increase in predicted genetic diversity from the mean results in a 24, 8, 8, 12, 12, and 13 percent increase in net productivity respectively; a significant change in terms of size.

Based on the results in Tables 4.4-4.6, no indications are present to support the notion that predicted genetic diversity affects net productivity in any way. We must take into account the possibility of irrelevant variables as well as the possibility of omitted variable bias; the original specification employed by Ashraf and Galor employed income per capita as the dependent variable. We must construct a specification which only controls for those variables influencing net productivity. Equation [4] must be streamlined. For the sake of ensuring that we fairly model the relationship, if any, between net productivity and genetic diversity, we present a new model which we estimate using OLS.

The ‘social infrastructure’ control variable is worth mentioning. Notice that a significant coefficient estimate of the ‘social infrastructure’ variable is reported 100 percent of the time. Of the 20 models in which the social infrastructure is controlled for, 95 percent of the models report a highly significant positive coefficient estimate. There could be a possibility that this statistically powerful variable dampens any significant effect of another variable upon those used to model economic developmental differences.

Tables A.3 and A.4 in the appendix display models in which we include the variable on respect and responsibility from Breuer and McDermott (2012). The coefficient estimate on genetic diversity is not the goal of these models. Notice that the inclusion of the respect and responsibility measurement causes the social infrastructure variable to lose its statistical significance in some of the models. It is likely that the changes are also a result of a lower sample size given the limited country observations on respect and responsibility compared with Tables 4.3-4.6. Models were also estimated replacing the social infrastructure variable with the variable measuring respect and responsibility (these models are not reported in this paper); little change occurred compared with Tables 4.4 and 4.6. The coefficient estimate on the summation of respect and responsibility were statistically significant in a majority of the models. The hump-shape relationship could still be observed as in Table 4.6 but the coefficient estimates were not statistically significant. In most cases, the results included in the appendix provide an avenue of justification; respect and responsibility and social infrastructure measurements are similar in the amount of variation they capture. These variables in addition to the variable measuring trust are further compared in Table 4.10.

TABLE 4.3: OLS Estimation using Cross-Sectional Data – Linear

	Dependent variable is log net productivity in 2000 CE					
	(1)	(2)	(3)	(4)	(5)	(6)
Predicted diversity (ancestry adjusted)	15.387** (7.009)	11.858* (6.629)	10.665* (6.103)	13.354** (5.566)	13.508** (6.067)	11.729* (6.116)
Log Neolithic transition timing (ancestry adjusted)	-.466 (.679)	-.049 (.480)	.014 (.488)	.011 (.543)	-.011 (.593)	.012 (.575)
Log percentage of arable land	-.271 (.172)	-.082 (.121)	-.083 (.114)	-.031 (.147)	-.037 (.152)	-.047 (.166)
Log absolute latitude	.250 (.189)	.326** (.161)	.244* (.136)	.245 (.167)	.232 (.198)	.173 (.205)
Social infrastructure		2.698*** (.703)	2.852*** (.630)	2.641*** (.787)	2.636*** (.823)	2.049*** (.775)
Ethnic fractionalization			-1.102* (.629)	-.930 (.596)	-.915 (.638)	-.863 (.568)
Years of schooling						
OPEC fixed effect	No	No	No	No	Yes	Yes
Legal origin fixed effects	No	No	No	Yes	Yes	Yes
Major religion shares	No	No	No	Yes	Yes	Yes
R^2	.73	.85	.87	.91	.91	.92
	N = 54					

*** p<.01, ** p<.05, * p<.0.1

Notes: Within this table, all regressions include Sub-Saharan Africa and continent fixed effects. Bootstrap standard errors (where replication = 1000), taking into account the use of generated explanatory variables, are presented in parentheses.

TABLE 4.4: OLS Estimation using Cross-Sectional Data - Quadratic

	Dependent variable is log net productivity in 2000 CE					
	(1)	(2)	(3)	(4)	(5)	(6)
Predicted diversity (ancestry adjusted)	153.624 (206.881)	73.63 (214.927)	11.223 (188.910)	-3.071 (193.670)	2.576 (227.654)	3.599 (243.957)
Predicted diversity square (ancestry adjusted)	-100.254 (150.184)	-44.781 (154.136)	-.404 (135.793)	11.943 (140.09)	7.940 (163.044)	5.905 (174.844)
Log Neolithic transition timing (ancestry adjusted)	-.305 (.752)	-.018 (.518)	.015 (.526)	-.010 (.616)	-.023 (.619)	.003 (.592)
Log percentage of arable land	-.267 (.173)	-.082 (.123)	-.083 (.115)	-.032 (.149)	-.037 (.166)	-.047 (.167)
Log absolute latitude	.227 (.213)	.315* (.187)	.244 (.151)	.247 (.180)	.234 (.218)	.175 (.215)
Social infrastructure		2.67*** (.707)	2.851*** (.636)	2.659*** (.778)	2.648*** (.911)	2.059** (.897)
Ethnic fractionalization			-1.101* (.648)	-.946 (.664)	-.926 (.667)	-.871 (.645)
Years of schooling						.129 (.090)
Optimal diversity	.76	.82		No	No	No
OPEC fixed effect	No	No	No	No	Yes	Yes
Legal origin fixed effects	No	No	No	Yes	Yes	Yes
Major religion shares	No	No	No	Yes	Yes	Yes
R^2	.74	.85	.87	.91	.91	.92
	N = 54					

*** p<.01, ** p<.05, * p<.0.1

Notes: Within this table, all regressions include Sub-Saharan Africa and continent fixed effects. Bootstrap standard errors (where replication = 1000), taking into account the use of generated explanatory variables, are presented in parentheses.

TABLE 4.5: OLS Estimation using Cross-Sectional Data (Hall and Jones) – Linear

	Dependent variable is log net productivity in 1988 CE					
	(1)	(2)	(3)	(4)	(5)	(6)
Predicted diversity (ancestry adjusted)	1.207 (3.088)	3.493 (2.713)	3.332 (2.788)	3.390 (2.527)	3.390 (2.584)	3.093 (2.706)
Log Neolithic transition timing (ancestry adjusted)	.372* (.212)	.375** (.152)	.381** (.166)	.340* (.182)	.340* (.187)	.295 (.204)
Log percentage of arable land	-.044 (.047)	-.006 (.040)	-.005 (.039)	.009 (.046)	.009 (.046)	.048 (.061)
Log absolute latitude	.139 (.096)	.101 (.081)	.100 (.082)	.131 (.087)	.131 (.087)	.093 (.094)
Social infrastructure		1.468*** (.279)	1.449*** (.301)	1.213*** (.322)	1.213*** (.313)	1.141*** (.410)
Ethnic fractionalization			-.053 (.264)	-.087 (.269)	-.087 (.265)	-.003 (.349)
Years of schooling						.021 (.041)
OPEC fixed effect	No	No	No	No	Yes	Yes
Legal origin fixed effects	No	No	No	Yes	Yes	Yes
Major religion shares	No	No	No	Yes	Yes	Yes
R^2	.50	.64	.64	.72	.72	.70
N	108	107	106	106	106	91

*** p<.01, ** p<.05, * p<.0.1

Notes: Within this table, all regressions include Sub-Saharan Africa and continent fixed effects. Bootstrap standard errors (where replication = 1000), taking into account the use of generated explanatory variables, are presented in parentheses.

TABLE 4.6: OLS Estimation using Cross-Sectional Data (Hall and Jones) – Quadratic

	Dependent variable is log net productivity in 1988 CE					
	(1)	(2)	(3)	(4)	(5)	(6)
Predicted diversity (ancestry adjusted)	112.661 (78.351)	30.647 (76.082)	32.259 (82.170)	49.919 (73.500)	49.919 (71.194)	55.843 (76.895)
Predicted diversity square (ancestry adjusted)	-79.156 (55.667)	-19.307 (54.085)	-20.574 (58.462)	-33.111 (52.508)	-33.111 (50.830)	-37.616 (55.088)
Log Neolithic transition timing (ancestry adjusted)	.381* (.196)	.377** (.160)	.382** (.160)	.347* (.190)	.347* (.181)	.312 (.219)
Log percentage of arable land	-.053 (.051)	-.009 (.040)	-.008 (.043)	.005 (.047)	.005 (.044)	.045 (.065)
Log absolute latitude	.128 (.098)	.099 (.084)	.098 (.086)	.128 (.088)	.128 (.088)	.095 (.104)
Social infrastructure		1.448*** (.303)	1.427*** (.317)	1.175*** (.350)	1.175*** (.329)	1.111*** (.421)
Ethnic fractionalization			-.046 (.258)	-.076 (.259)	-.076 (.256)	.013 (.313)
Years of schooling						.018 (.042)
Optimal diversity	.71	.79	.78	.75	.75	.74
OPEC fixed effect	No	No	No	No	Yes	Yes
Legal origin fixed effects	No	No	No	Yes	Yes	Yes
Major religion shares	No	No	No	Yes	Yes	Yes
R^2	.51	.64	.64	.72	.72	.70
N	108	107	106	106	106	91

*** p<.01, ** p<.05, * p<.0.1

Notes: Within this table, all regressions include Sub-Saharan Africa and continent fixed effects. Bootstrap standard errors (where replication = 1000), taking into account the use of generated explanatory variables, are presented in parentheses.

4.1.2 Net Productivity and Predicted Genetic Diversity

Reconstructed Model

In re-estimating the relationship between predicted diversity and net productivity, the following model is tested.

$$ln\, B_i = \delta_0 + \delta_1 \hat{G}_i + \delta_2 \widehat{G_i^2} + \dddot{\delta}_3 \mathrm{I}_i + \mu_i \qquad [7]$$

where I_i represents the vector for institutional, geographical, and social variables. Table 4.8 displays the coefficient estimates with respect to equation [7], excluding the quadratic possibility. The significance of the coefficient estimate on predicted diversity becomes more prevalent as we correct equation [4] to take into the account the possibility of irrelevant variables. Models 1-3 in Table 4.8 generate significant and positive coefficient estimates of predicted diversity at the 5 percent level; models 4-5 also display significant and positive coefficient estimates at the 10 percent level. The five models roughly explain just as much of the variation in net productivity as did the original models constructed based on the Ashraf and Galor specification. Specifically, a 1 percent increase in predicted diversity from the mean results in a 6.4 percent increase in net productivity based on model 1. Based on models 2-5, a 1 percent increase in predicted diversity from the mean results in a 6.7, 7.3, 7.7, and a 7.8 percent increase in net productivity respectively; these magnitudes are relatively large.

Focusing on the sign of the coefficient estimates, all models generate a positive effect of predicted diversity on net productivity. The significant coefficient estimates on the controlled variables all display their anticipated signs. The variables ‘absolute latitude' and 'social infrastructure' are both positive and significant. Ethnic fractionalization displays a significant negative effect on net productivity as anticipated. The remaining variables, years of schooling, democracy, and executive constraint all generate statistically insignificant coefficient estimates thus little can be inferred by their results. Neolithic transition timing, the dummy variable for OPEC membership, and the major religion share variables were all excluded from the reconstructed model; these variables fail to prove significant in terms of a direct effect upon net productivity in the year 2000. It is important to note that even if one controls for the Neolithic transition timing variable, the coefficient estimate upon predicted diversity is still statistically significant. Based on Table 4.8, empirical results support the notion that predicted diversity positively affects net productivity upon a given population; the characteristics of equation [1] are validated.

The primary relationship we are concerned with would be the quadratic relationship between net productivity and genetic diversity. Our previous models have been inconsistent regarding what we can infer about that very relationship, if it indeed exists. Applying what we understand about genetic diversity from Chapter 2 will allow us to reasonably construct a model with accurate control variables. First, Ashraf and Galor argue that genetic diversity plays a role in influencing such things as knowledge creation or educational attainment to a certain degree and social behaviors, particularly that of cooperation and trust. The 'social infrastructure' variable has proven to be a highly significant control in the models presented so far. Social infrastructure likely is creating a 'crowding out' effect on the genetic diversity variable; genetic diversity is an underlying cause of variations in social infrastructure thus we must separate this effect with the removal of social infrastructure. Overall, once social infrastructure is introduced, it acts as an ‘answer’ to diminish any impacts observed through genetic diversity.

Second, when taking into account the quadratic relationship, we are allowing the possibility for there to exist some degree of separation among groups as a result of non-cooperation or lack of trust. Ethnic fractionalization dampens the effects caused by deteriorations in social capital; it too can be removed. The variable measuring years of schooling can likely ‘pick up’ the byproducts associated with increasing genetic diversity levels, particularly that of knowledge creation; it too can be removed. Lastly, the variables measuring democracy and executive constraints are highly correlated with each other; including both likely leads to errors within our coefficient estimate variances. One can reasonably argue that both democracy and executive constraints have little influence in terms of affecting net productivity. Once religion shares and legal origins are controlled for, it is likely that democracy and constraint on the executive will add little to the fit of the model; so, both are removed.

Lastly, the model takes into consideration many of the theoretical assumptions made about the genetic diversity variable. Many of the ‘key factors’ necessary to observe changes in productivity are already taken into account with genetic diversity as presented by Ashraf and Galor (2013); education, technology, behaviors etc. We conform to these theoretical arguments when streamlining our model; the results do support the notion that a hump-shape relationship exists between net productivity and genetic diversity. Table 4.9 presents the highly significant hump-shape relationship for models 1-3. Model 4, once all of our primary controls are factored in, is significant at the 10 percent level. Notice our optimal degrees of genetic diversity are each .70 throughout the models, consistent with that of Ashraf and Galor (2013). The R-squared level reinforces the strength of these few key variables as well; model 4 explains nearly 68 percent of the variation in net productivity. The variables all take on long-term characteristics, if not permanent, relative to previous variables used; these variables compliment the prehistoric nature of the genetic diversity variable.

4.1.3 Productivity and Social Capital

Respect and Responsibility

TABLE 4.7: Summary Statistics for Table 4.10

Variable	Mean	Std. Dev.	Min	Max
Predicted diversity (ancestry adjusted)	0.723	0.026	0.643	0.765
Log of income per capita in 2000	9.174	0.926	6.964	10.445
Social infrastructure	0.561	0.268	0.113	0.973
Ethnic fractionalization	0.326	0.243	0.002	0.930
Years of schooling	5.728	2.553	1.572	10.862
Democracy	5.435	3.780	0.000	10.000
Executive constraint	4.806	1.890	1.000	7.000
Respect and Responsibility (Average)	0.720	0.093	0.516	0.895
Respect and Responsibility (Sum)	1.440	0.185	1.033	1.790

Table 4.10 displays the coefficient estimates of predicted diversity on the respect and responsibility measurements obtained from Breuer and McDermott (2012) alongside the “social infrastructure” and “trust” variables. Based on the results, the respect and responsibility measurements are the only measurements of the three which are negatively associated with social capital; the lack of statistical significance in model 1 is shown.

Table 4.10 provides empirical support regarding our opinion that respect and responsibility, in combination, are the better variables to serve as a proxy for social capital relative to the degree of trust in Ashraf and Galors' study. This empirical study assumes that social capital, the variable discussed in Ashraf and Galor (2013), is primarily affected by the level of respect and prevalence of responsibility in a given population. In order to observe the effect of genetic diversity on an aggregated measure of respect and responsibility, we employ the average of each variable as well as the combined value of the respect and responsibility variables; these dependent variables serve as proxy variables for social capital. Table 4.10 tests the notion presented in Ashraf and Galor that increasing levels of genetic diversity negatively affect cooperation among a given population. The negative effect helps to reinforce the results in Ashraf and Galor regarding the hump-shape relationship generated between per capita income and genetic diversity.

Model 1 in Table 4.10 tests the relationship between predicted diversity and the summation of respect and responsibility across country. A statistically significant and negative relationship exists between diversity and the proxy variable for social capital at the 5 percent level. Roughly 69 percent of the variation in respect and responsibility across countries are explained by model 1. Specifically, a 1 percent increase in predicted diversity from the mean results in a .011 unit decrease in the total respect and responsibility measurement; roughly a .76 percent decrease in the total respect and responsibility measurement from its mean. The magnitude is moderate regarding the size of the marginal effect of diversity upon the respect and responsibility measurements.

Model 2 in Table 4.10 tests the relationship between predicted diversity and the average of respect and responsibility across country. A statistically significant and negative relationship exists between diversity and the proxy variable for social capital at the 5 percent level. Roughly 69 percent of the variation in respect and responsibility across countries are explained by model 2. Specifically, a 1 percent increase in predicted diversity from the mean results in a .005 unit decrease in the average respect and responsibility level; roughly a .34 percent decrease in the average of respect and responsibility from its mean. The magnitude is not as dramatic regarding the impact on respect and responsibility from a change in predicted diversity. These two models support the argument presented by Ashraf and Galor regarding genetic diversity's effect upon social capital.

4.2 Critiques

Based on our findings, there is evidence to suggest that predicted genetic diversity plays a central role in explaining variations in net productivity across countries. With genetic diversity's effect upon net productivity, it is clear how the hump-shape relationship between income per capita and diversity is generated within the Ashraf and Galor study. Ashraf and Galor seek to argue that net technological productivity affects output per worker (or labor augmented productivity) which in turn will affect output per capita. This study proves that net technological productivity does affect labor-augmented productivity through genetic diversity.

Although the study provides support for the hump-shape relationship associated between income per capita and genetic diversity, there exists little degree of robustness once we test net productivity against the genetic diversity measurement. As controls are added to the models presented in Table 4.9, the significance is no longer observed. We hypothesize further, suggesting that although the Ashraf and Galor study is further justified, its impact regarding current economic development is limited. Limitations on the genetic diversity variable refute suggestions made by Ashraf and Galor that its effect is long-lasting.

TABLE 4.8: Reconstructed OLS Estimation using Cross-Sectional Data – Linear

	Dependent variable is log net productivity in 2000 CE				
	(1)	(2)	(3)	(4)	(5)
Predicted diversity (ancestry adjusted)	8.944** (3.898)	9.437** (4.160)	10.258** (3.878)	10.720*** (3.922)	10.965*** (3.984)
Log absolute latitude	.270** (.112)	.284** (.114)	.232** (.112)	.234** (.112)	.252** (.114)
Social infrastructure	2.994*** (.484)	3.126*** (.536)	1.910*** (.589)	1.962*** (.616)	2.078*** (.626)
Ethnic fractionalization	-1.371*** (.459)	-1.377*** (.465)	-1.186*** (.372)	-1.212*** (.389)	-1.294*** (.405)
Years of schooling		-.027 (.052)	.071 (.059)	.077 (.062)	.071 (.062)
Democracy				-.016 (.052)	-.122 (.146)
Executive Constraint					.194 (.253)
Legal origin fixed effects	No	No	Yes	Yes	Yes
R^2	.86	.86	.90	.90	.90
	N = 54				

*** p<.01, ** p<.05, * p<.0.1

Notes: Within this table, all regressions include continent fixed effects. Robust standard errors are presented in parentheses.

TABLE 4.9: Reconstructed OLS Estimation using Cross-Sectional Data – Quadratic

	Dependent variable is log net productivity in 2000 CE			
	(1)	(2)	(3)	(4)
Predicted diversity (ancestry adjusted)	636.111*** (172.487)	565.028*** (156.076)	600.441*** (173.036)	272.221* (154.555)
Predicted diversity squared (ancestry adjusted)	-452.743*** (123.990)	-402.003*** (112.048)	-423.640*** (124.87)	-192.587* (111.71)
Log absolute latitude				.733* (.149)
Major Religion Shares	No	No	Yes	Yes
Legal origin fixed effects	No	Yes	Yes	Yes
R^2	.15	.35	.49	.68
Optimal diversity	.70	.70	.70	.70
N = 54				

*** p<.01, ** p<.05, * p<.0.1

Notes: Within this table, all regressions include continent fixed effects. Robust standard errors are presented in parentheses.

We cannot expect to learn a great deal from prehistoric factors such as genetic diversity in the contemporary era; the lack of robustness informs us of this. One reason the Ashraf and Galor study could have avoided robustness errors as more controls were added could be because of genetic diversity's immense role as a primary underlying effect within many different variables. Genetic diversity likely picks up a great degree of effects from omitted variables in the Ashraf and Galor study. This does not suggest that their findings are not correct to a degree, what this suggests is that genetic diversity has a limited effect upon the economic development conditions of any given country.

It is likely that the effect of genetic diversity on the economic development potential of a population was more direct prior to advancements in technological infrastructure which weakened many of the limitations resulting from a population's genetic make-up. As time progressed, advancements in the world's technological capabilities alongside the onset of globalization acted to expand the limitations observed in populations experiencing too much diversity or too little diversity. This is not to say that genetic diversity cannot explain the variation in cross-country economic levels but that this explanation is no longer directly observed; there are solutions present which can undo the trajectory that countries were placed on as a result of genetic diversity; utilizing these solutions effectively or not does not enhance the significance of genetic diversity in terms of its modern importance. Based on the results, contemporary factors are more suitable in explaining variations regarding the economic positioning of countries.

Aside from the empirical critiques, one could argue that Ashraf and Galor fail to provide any reasonable solutions for solving the income inequality observed among countries. Perhaps, the difficulty in constructing solutions to counteract prehistoric factors is the very reason why little regarding this matter is mentioned. Assuming genetic diversity acts as a long-lasting direct factor associated with the contemporary economic divergence of countries, Ashraf and Galor indirectly argue that not much can be done about this inequality. This study takes a more optimistic perspective regarding the opportunities available to close the economic gap observed between countries.

TABLE 4.10: Respect and Responsibility

	(1) Social Infrastructure	(2) Degree of Interpersonal Trust	(3) Respect and Responsibility (Sum)	(4) Respect and Responsibility (Average)
Predicted diversity (ancestry adjusted)	- .342	.331	-1.674**	-.837**
	(.845)	(1.069)	(.633)	(.316)
Log of income per capita in 2000	.235***	.047	.068	.034
	(.044)	(.068)	(.048)	(.024)
Ethnic fractionalization	.164	.070	.120	.060
	(.097)	(.139)	(.091)	(.045)
Years of schooling	-.003	-.013	.012	.006
	(.019)	(.024)	(.018)	(.009)
Democracy	.041*	.001	.025	.012
	(.021)	(.030)	(.026)	(.013)
Executive constraint	-.063	.000	-.029	-.014
	(.042)	(.061)	(.054)	(.027)
Legal origin	Yes	Yes	Yes	Yes
R2	.83	.33	.69	.69
N	43	43	43	43

*** p<.01, ** p<.05, * p<.0.1

Notes: Robust standard errors are presented in parentheses.

CHAPTER 5

CONCLUSION

This study sought to test the validity of the Ashraf and Galor “Out of Africa” hypothesis. The authors asserted that the aggregate production process was affected by two counteractive forces, a negative effect resulting from genetic diversity’s effect upon social capital and a positive effect resulting from genetic diversity’s effect upon technological productivity. As a result of their hypothesis, a hump-shape relationship should be observed between per capita income and predicted genetic diversity; their results reveal this to be true. This study sought to test the underlying assumptions, which they claim justify the significant hump shape relationship observed. Specifically, if their analysis is true, a hump-shape relationship should also be observed between predicted genetic diversity and net productivity.

Based on the Ashraf and Galor study, increases in diversity first, lead to knowledge creation which positively affects the technological production process within the given population; this positive effect diminishes as diversity levels continue to increase. As diversity levels continue to rise, non-cooperation resulting from deterioration in social capital leads to negative effects upon the production process. This negative effect results in the foregone gross productivity modeled in equation [2]; net productivity thus is the observed outcome. Their study continues to assert that these counteractive effects directly affect output per worker (or labor-augmented productivity). This theoretical argument is used to justify their hypothesis concerning a hump-shape relationship between per capita output and genetic diversity.

Net productivity should be related to predicted diversity in a hump-shape fashion in order for their exact theory to be credible. Net productivity is gross productivity less the negative effects of social capital as diversity levels rise within a given population. In an effort to test the underlying assumptions, we constructed a measure of net productivity following the methodology of Hall and Jones (1998). Once we constructed our net productivity measurement, we then used this measurement within our regression as well as appropriate explanatory variables to test the validity of Ashraf and Galors' results. Several models were taken into account in an effort to provide flexibility to the functional form of the model.

With respect to the empirical models, we focused upon the relationship between net productivity and genetic diversity utilizing both the constructed net productivity measurements as well as those of Hall and Jones (1998). The models were further streamlined to take into account irrelevant and/or omitted variables. The overall results did provide reasonable evidence to suggest a hump-shape relationship exist between net productivity and genetic diversity. Following the estimation of net productivity to predicted genetic diversity, a model was constructed to test the effect, if any, that diversity had upon the social capital of a given population. We did observe a significant negative relationship between the social capital variables, respect and responsibility, and genetic diversity.

Overall, we conclude that although a hump-shape relationship is observed between net productivity and genetic diversity, the lack of robustness provides evidence to suggest that the genetic diversity variable has its limitations in terms of explaining the variation observed in net productivity across countries. It is likely that genetic diversity acts as an underlying affect for a greater amount of variables affecting income per capita relative to net productivity; this could explain the significant hump-shape relationships observed in Ashraf and Galors' models when testing robustness. Additionally, it would be more beneficial to focus on explanatory variables which affect the economic development potential of countries while also possessing the ability to be altered; specifically as a result of policy initiatives.

REFERENCES

Alesina, Alberto, and Arnaud Devleeschauwer, William Easterly, Sergio Kurlat, and Romain Wacziarg. 2003. "Fractionalization." *Journal of Economic Growth*, 8(2): 155-194.

Ashraf, Quamrul and Oded Galor. "The 'Out of Africa' Hypothesis, Human Genetic Diversity, and Comparative Economic Development." *American Economic Review* 103 (2013): 1-46. Web.

Ashraf, Quamrul and Oded Galor. "Isolation and Development." (2009): 1-12. Web.

Berlemann, Michael and Jan-Erik Wesselhöft. "Estimating Aggregate Capital Stocks Using the Perpetual Inventory Method." 125 (2012): 1-42. Web.

Breuer, Janice and John McDermott, "Respect, Responsibility, and Development." (2012): 31. Forthcoming, *Journal of Development Economics*, 2014.

Cavalli-Sforza, Luca L. "The Human Genome Diversity Project: past, present and future." <u>Nature Reviews</u> Apr. 2005: 333-340. Web.

Guedes, Jade d'Alpoim, et al. "Is Poverty in Our Genes?" *Current Anthropology* 54 (2013): 71-79. Web.

Hall, Robert E. and Charles I. Jones. "Why Do Some Countries Produce So Much More Output per Worker than Others?" (1998): 1-14. Web.

La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer, and Robert W. Vishny. 1999. "The Quality of Government." *Journal of Law, Economics, and Organization*, 15(1): 222-279.

McEvedy, Colin, and Richard Jones. 1978. *Atlas of World Population History*. New York, NY: Penguin Books Ltd.

McGue, Matt and Thomas J. Bouchard, Jr. "Genetic and Environmental Influences on Human Behavioral Differences." *Annual Review of Neuroscience* 24 (1998): 1-24. Web.

Montalvo, Jose G and Marta Reynal-Querol. "Ethnic Diversity and Economic Development." *Journal of Development Economics* 76 (2005): 293-323. Web.

Putterman, Louis. 2008. “Agriculture, Diffusion, and Development: Ripple Effects of the Neolithic Revolution.” *Economica*, 75(300): 729-748.

Ramachandran, Sohini, et al. “Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa.” PNAS Nov. 1 2005: 15942-15947. Web.

Rosenberg, Noah A, et al. “Genetic Structure of Human Populations.” *Science* 298 (2002): 2381-2385. Web.

APPENDIX A

OLS Estimations Using Respect and Responsibility

TABLE: A.1 Summary Statistics for Table A.3

Variable	Mean	Std. Dev.	Min	Max
Predicted diversity (ancestry adjusted)	0.720	0.025	0.643	0.765
Predicted diversity squared (ancestry adjusted)	0.519	0.036	0.414	0.586
Log Neolithic transition timing	8.704	0.268	8.161	9.173
Log percentage of arable land	2.876	0.841	1.061	4.129
Log absolute latitude	3.402	0.871	0.000	4.159
Social infrastructure	0.603	0.269	0.156	0.973
Ethnic fractionalization	0.308	0.251	0.012	0.930
Years of schooling	6.029	2.637	1.572	10.862
Respect and Responsibility (Sum)	1.457	0.194	1.033	1.790
Log net productivity	-1.346	1.314	-4.112	0.019

TABLE: A.2 Summary Statistics for Table A.4

Variable	Mean	Std. Dev.	Min	Max
Predicted diversity (ancestry adjusted)	0.721	0.026	0.643	0.756
Predicted diversity squared (ancestry adjusted)	0.520	0.036	0.414	0.572
Social Infrastructure	0.584	0.259	0.156	0.973
Ethnic fractionalization	0.302	0.233	0.002	0.752
Log Neolithic transition timing	8.707	0.367	7.300	9.250
Log absolute latitude	3.411	0.738	0.312	4.159
Log percentage of arable land	2.635	0.972	0.399	4.129
Years of Schooling	5.874	2.448	1.572	10.862
Respect and Responsibility (Sum)	1.447	0.182	1.033	1.790
Log of net productivity	-0.517	0.606	-2.244	0.188

TABLE A.3: OLS Estimation using Cross-Sectional Data – Quadratic

	Dependent variable is log net productivity in 2000 CE				
	(1)	(2)	(3)	(4)	(5)
Predicted diversity (ancestry adjusted)	-92.694 (674.133)	-192.386 (682.636)	-37.354 (4873.495)	-699.053 (2539.707)	-566.232 (4094.911)
Predicted diversity square (ancestry adjusted)	70.211 (479.309)	143.658 (482.562)	33.155 (3468.133)	493.703 (1797.759)	393.012 (2794.461)
Log Neolithic transition timing (ancestry adjusted)	-.528 (.928)	-.524 (.884)	-.860 (7.355)	-.480 (3.954)	.128 (11.373)
Log percentage of arable land	.005 (.269)	-.001 (.298)	.088 (2.146)	.315 (1.440)	.276 (2.995)
Log absolute latitude	.414 (.547)	.189 (.634)	.556 (4.604)	2.278 (7.518)	2.283 (21.125)
Social infrastructure	3.093*** (1.152)	3.363*** (1.219)	3.007 (20.178)	1.832 (11.096)	1.119 (12.458)
Respect & Responsibility (Sum)	.274 (1.069)	.023 (1.083)	.088 (11.324)	-.870 (7.893)	-.991 (16.605)
Ethnic fractionalization		-1.118 (.958)	-.828 (5.602)	-.793 (3.683)	-.804 (11.142)
Years of schooling					.143 (.537)
OPEC fixed effect	No	No	No	Yes	Yes
Legal origin fixed effects	No	No	Yes	Yes	Yes
Major religion shares	No	No	Yes	Yes	Yes
R^2	.85	.87	.94	.96	.97

N = 32

*** p<.01, ** p<.05, * p<.0.1

Notes: Robust standard errors are presented in parentheses.

TABLE A.4: OLS Estimation using Cross-Sectional Data (Hall and Jones) – Quadratic

	Dependent variable is log net productivity in 2000 CE				
	(1)	(2)	(3)	(4)	(5)
Predicted diversity (ancestry adjusted)	8.843 (366.587)	20.825 (351.582)	69.646 (312.463)	69.646 (282.617)	88.881 (458.914)
Predicted diversity square (ancestry adjusted)	-.937 (258.443)	-9.707 (248.662)	-48.506 (220.563)	-48.506 (200.439)	-62.682 (330.629)
Log Neolithic transition timing (ancestry adjusted)	.157 (.699)	.151 (.726)	.472 (.674)	.472 (.674)	.501 (1.235)
Log percentage of arable land	.023 (.172)	.022 (.192)	.010 (.195)	.010 (.213)	.008 (.276)
Log absolute latitude	-.066 (.288)	-.052 (.318)	-.029 (.370)	-.029 (.345)	-.044 (.344)
Social infrastructure	1.600 (.742)	1.582** (.797)	.641 (.962)	.641 (1.052)	.560 (1.173)
Respect & Responsibility (Sum)	.186 (.625)	.216 (.667)	1.355 (1.018)	1.355 (1.041)	1.317 (2.068)
Ethnic fractionalization		.159 (.631)	-.141 (.636)	-.141 (.691)	-.145 (1.717)
Years of schooling					.023 (.240)
OPEC fixed effect	No	No	No	Yes	Yes
Legal origin fixed effects	No	No	Yes	Yes	Yes
Major religion shares	No	No	Yes	Yes	Yes
R^2	.52	.53	.85	.85	.85

N = 41

*** p<.01, ** p<.05, * p<.0.1

Notes: Robust standard errors are presented in parentheses.