UNIVERSIDADE DA BEIRA INTERIOR
Ciências Sociais e Humanas

Essays on Economic Development, Human Capital, Technology and Inequality

Marcelo Serra Santos

Tese para obtenção do Grau de Doutor em
Economia
(3º ciclo de estudos)

Orientador: Prof. Doutor Tiago Miguel Guterres Neves Sequeira
Co-orientador: Prof.ª Doutora Alexandra Maria do Nascimento Ferreira Lopes

Covilhã, 2015

Agradecimentos

Na realização de uma tese, embora seja um trabalho de foro individual, envolve sempre um grupo de trabalho mais alargado existindo assim algumas individualidades que merecem o seu destaque não só pelo contributo direto como indireto. Por tudo isto expresso os meus agradecimentos aos seguintes:

Ao Professor Doutor Tiago Sequeira, orientador desta tese, pela sua disponibilidade, interesse, simpatia e ajudas constantes, pela partilha do saber e seus incentivos.

À Professora Doutora Alexandra Lopes, co-orientadora desta tese, pela sua dedicação, sabedoria, preocupação e rigor que contribuíram, certamente, para o enriquecimento deste trabalho.

Ao Centro de Estudos e Formação Avançada em Gestão e Economia (CEFAGE) pelas condições disponibilizadas, através dos vários workhops realizados, contribuíndo para o diálogo, partilha de conhecimento e incentivo por parte das várias personalidades envolvidas. Entre estes merece o destaque a Professora Doutora Andreia Dionísio, Professora Doutora Esmeralda Ramalho e Professor Doutor Miguel Rocha de Sousa.

Aos organizadores e restantes intervenientes da conferência realizada pela *Eurasia Business and Economics Society* (EBES) que possibilitou a partilha de conhecimentos e a análise crítica contribuindo para melhorar o presente estudo.

À Universidade Nova School of Business and Economics, em particular à Professora Doutora Ana Balcão Reis, por oferecer a oportunidade de esclarecer questões importantes à realização deste trabalho com o Professor Doutor Diego Comin e ainda por poder assistir à sua conferência. Como não poderia deixar de ser, o meu agradecimento ao próprio Professor Doutor Diego Comin que, com muita simpatia, gentileza e saber, esclareceu todas as questões colocadas e, contribuíndo assim, ao desenvolvimento deste estudo.

Aos meus amigos Luís Marques, Gonçalo Marques, Bruma Almeida, Tiago Batista, Fabiana Vicente, Dan Costin, Diana França e Guillaume Borges pelos seus incentivos e amizade.

Aos meus pais, irmãos e restantes familiares pelos incentivos constantes.

À Marisa, minha companheira, por tudo o que representa para mim, pelo seu amor e apoio incondicional.

Resumo alargado

Esta tese contém cinco ensaios sobre o desenvolvimento económico, o capital humano, a tecnologia e a desigualdade. A literatura recente é centrada na explicação do crescimento através das instituições. De facto, a qualidade das instituições tem sido associada com a *performance* do crescimento económico dos países e também dos seus níveis de desenvolvimento. Determinantes como a cultura, geografia, língua, religião, composição étnica e genética das populações, e o poder colonial de um determinado país (conjunto a que se designou como factores profundamente enraizados na história) têm sido estudados bem como as suas ligações com as instituições atuais, a qualidade dos governos e o comércio. Além disso, a origem da desigualdade de rendimento é um dos assuntos primordiais para os economistas porque, para além de ser difícil encontrar as suas raízes, ela é também sensível a algumas mudanças que são caracteristicas da evolução recente da humanidade, como é o caso da educação ou do crescimento. *Os três primeiros ensaios desta dissertação estão inseridos na literatura relacionada com os factores de desenvolvimento profundamente enraizados na história. Os dois últimos ensaios inserem-se na literatura que explica a desigualdade de rendimentos.*

Os estudos empíricos avaliaram então os determinantes profundamente enraizados dos vários caminhos do desenvolvimento, onde a *cultura*, a religião, as instituições, a genética e as características biogeográficas foram alguns dos aspetos testados. Guiso *et al.* (2009) tentaram capturar os efeitos da cultura nas trocas económicas. Eles descobriram que as variáveis históricas e culturais, como a semelhança religiosa, confiança, raízes linguísticas comuns ou a mesma origem legal, afectam as relações de confiança entre dois países e também as diferenças de confiança tendem a afectar o comércio, os investimentos e o investimento directo estrangeiro. Num diferente ponto de vista, Lada (2013) estudou os efeitos da similaridade cultural (usando algumas medidas diferentes como a proximidade racial, religião e civilização) e o seu efeito nos conflitos entre 1816 e 2009, concluindo que a similaridade cultural provoca mais hostilidade e aumenta a propensão à guerra gerando portanto incerteza e reduzindo o crescimento económico. Um dos mais importantes estudos deste campo é Ashraf e Galor (2011a), que tentou obter relações históricas de assimilação cultural e difusão cultural (medido pela diversidade cultural e pelo índice de isolamento) com os diferentes padrões de desenvolvimento económico em todo o mundo. Empiricamente, com os métodos OLS (método dos mínimos quadrados) e IV (método com variáveis instrumentais), eles concluíram, com o uso de uma amostra com 60 países, que o isolamento geográfico na era pré-industrial, teve um impacto negativo sobre a diversidade cultural na era moderna, medida pelo índice de diversidade proveniente dos dados do World Value Survey (WVS), mas este efeito de isolamento sobre a diversidade cultural tem sofrido uma diminuição ao longo do século XX. Os seus resultados indicam que o aumento de um desvio padrão no isolamento geográfico provoca um decréscimo na diversidade cultural em 0,48 desvios-padrão. O isolamento geográfico teve um impacto positivo no desenvolvimento económico na era agrícola, em particular, onde existiu um efeito positivo do índice de isolamento no logaritmo de densidade populacional nos anos 1, 1000 e 1500 DC (Depois de Cristo). Além disso, o isolamento geográfico teve um impacto negativo sobre a rendimento *per capita* na era industrial, especificamente o índice de isolamento teve um efeito estatisticamente significativo e robusto no logaritmo de rendimento *per capita* em 1960, quando a maioria das nações estavam em processo de industrialização. A diversidade cultural teve um impacto positivo no desenvolvimento económico no processo de industrialização, especificamente houve um efeito positivo do índice de diversidade no logaritmo de rendimento *per capita* em 1960. Além disso, a *religião* pode ser um factor importante para o crescimento e desenvolvimento económico, uma vez que fez expandir as civilizações em todo mundo e aguentar-se historicamente mas também gerou conflitos, em particular entre as sociedades muçulmana e cristã, que levou ao fraccionamento religioso dentro das fronteiras nacionais. De acordo com Iyigun (2011) esta falta de cooperação é realmente adversa para o desenvolvimento económico. La Porta *et al.* (1999), após a sua investigação acerca dos determinantes de qualidade dos governos de 152 países, usando o método OLS, concluiu que os países em que a população é predominantemente protestante têm melhores governos do que paises onde as religiões predominantes são a católica ou a muçulmana. Estudos teóricos foram realmente importantes para o estabelecimento de algumas reflexões sobre as *instituições*. Tabellini (2008) acrescenta que as características das instituições políticas do passado deixaram uma marca nas atitudes e valores actuais, por exemplo, as regiões onde os governos anárquicos exploraram os seus cidadãos acabaram por desenvolver uma cultura de desconfiança. Acemoglu *et al.* (2001) tentou provar se as diferenças na experiência colonial levam a diferenças nas instituições usando duas amostras diferentes. Chegaram à conclusão de que há uma correlação entre a colonização europeia e as primeiras instituições e também correlação entre as primeiras instituições e as instituições de hoje, por isso a colonização construiu as suas instituições que se mantiveram ao longo do tempo, mantendo algumas das suas características ainda hoje. Hall e Jones (1999) tentaram perceber porque alguns países investem mais em capital físico e humano do que outros e porque alguns são mais produtivos do que outros. O estudo incluiu 127 países e os seus principais resultados mostram que as grandes diferenças de rendimento entre países são causadas por diferenças na infra-estrutura social (ligado às instituições que protegem os direitos de propriedade), que implicam grandes diferenças de acumulação de capital, níveis de escolaridade e produtividade. Ahlerup *et al.* (2009) testaram um modelo teórico, por meio de regressões de crescimento de *cross-country*, onde se constatou que o efeito marginal de capital social (confiança interpessoal) diminui com a força das instituições do país. A diversidade populacional pode ser útil para a produtividade, pois está completamente ligada à teoria da complementaridade de competências. Na verdade, diferentes trabalhadores têm diferentes maneiras para reagir a certos problemas no trabalho, o que afecta as suas capacidades produtivas. De facto, a diversidade pode causar opiniões divergentes, crenças e preferências que podem gerar falta de confiança, barreiras de comunicação ou mesmo conflitos. O trabalho de Ashraf e Galor (2013), com o objectivo de encontrar provas de como a *diversidade genética* é um determinante importante para o desenvolvimento económico, é um dos mais influentes neste campo. Ele usou a densidade populacional e o rendimento *per capita* como variáveis dependentes, a diversidade genética como a principal variável explicativa e diversas variáveis de controlo e variáveis instrumentais que tornaram os seus resultados mais robustos. Os autores verificaram que um êxodo do *Homo sapiens* para fora de África afectou a diversidade genética e teve historicamente um efeito duradouro sobre o padrão de desenvolvimento económico comparativo. A diversidade genética tem um efeito de U invertido sobre o desenvolvimento económico na era pré-colonial. Assim, a diversidade genética não tem um efeito monótono no rendimento *per capita* no mundo moderno. Baixos e altos níveis de diversidade são desvantajosos para o desenvolvimento, mas, por outro lado, níveis intermediários de diversidade podem causar mais desenvolvimento. Procurando efeitos de variáveis geográficas, históricas e culturais sobre o rendimento e produtividade, Spolaore e Wacziarg (2013a), afirmaram que a distância genética e os seus efeitos sobre as diferenças de rendimento atingiram o auge na segunda metade do século XIX e que começaram a diminuir na segunda metade do século XX. Também relatam um indirecto e persistente efeito no tempo dos factores biogeográficos pré-históricos que também é referido por Ashraf e Galor (2011b). Já

Easterly e Levine (1997) determinaram que o fraco crescimento e rendimento de África estão associados com a instabilidade política, os mercados cambiais distorcidos, a baixa escolaridade, os sistemas financeiros subdesenvolvidos e as infra-estruturas insuficientes. Tendo em conta um estudo mais diversificado, que inclui 195 países, entre 1990 e 2000, usando a diversidade do local de nascimento, uma variável nunca antes usada, Alesina *et al.* (2012) encontraram um efeito positivo da diversidade local de nascimentos de cidadãos nacionais e dos imigrantes sobre a produtividade, relação que é mais forte em países altamente produtivos. Além disso, há também evidência de efeitos positivos da diversidade do local de nascimento para os trabalhadores não qualificados sobre a produtividade. Para Comin *et al.* (2012), a distância espacial é determinante para a difusão: países que estão longe de líderes de adopção têm uma adopção lenta; contudo, os efeitos espaciais desaparecem ao longo do tempo. *Uma das conclusões desta revisão crítica de literatura é que a maioria dos artigos relacionados com os determinantes profundamente enraizados do desenvolvimento têm analisado empiricamente o seu efeito nas proxies de desenvolvimento, tais como a densidade populacional ou o rendimento per capita. Apenas uma minoria tem estudado o efeito desses determinantes historicamente enraizadas no capital humano (por exemplo Hall e Jones, 1999) ou no progresso tecnológico (por exemplo, Comin et. al., 2013). Os três primeiros ensaios desta tese contribuem para preencher esta lacuna. Em particular, estudamos o efeito da diversidade genética em diversas medidas de capital humano (1º ensaio), em diversas medidas de adopção tecnológica (2º ensaio) e analisamos os determinantes das invenções tecnológicas durante o processo da revolução industrial (3º ensaio). Esta tese destaca uma relação em forma de U invertido robusta entre a diversidade genética e os actuais níveis de capital humano e entre a diversidade genética e o nível pré-colonial de adopção tecnológica. Além disso, também confirma a importância dos efeitos de escala e da distância para EUA e Reino Unido como importantes impulsionadores para as inovações durante o processo da revolução industrial.*

A abordagem teórica para a desigualdade e nível de rendimento *per capita* teve inicio claro através da apresentação da curva de Kuznets (1955), que desenvolveu um modelo que assenta na realocação de pessoas e recursos provenientes da agricultura para a indústria. Mais tarde, essa relação começou por ser descrita como um U invertido; ou seja, a desigualdade aumenta até atingir um ponto, a partir do qual cai de novo, à medida que a economia se desenvolve. Allen (2003) testou seis principais explicações para o reverso no progresso da zona Mediterrânica contra o Reino Unido e os Países Baixos no início da Europa moderna - população, terra, papel do império, papel do governo, tecnologia e alfabetização; e concluiu que o papel do império (através do comércio internacional) foi o factor mais importante para o reverso do desenvolvimento e do aumento da desigualdade nos países da zona Mediterrânica. Allen (2009b), com uma revisão da literatura sobre a Revolução Industrial e o seu efeito sobre a desigualdade, percebeu que nas primeiras quatro décadas do século XIX a desigualdade aumentou consideravelmente, a taxa de lucro duplicou, o salário médio estagnou e o crescimento não pôde ser explicado sem a desigualdade, porque estes se influenciaram mutuamente. Esta relação parece ser mais forte quando os países são divididos em dois grupos: pobres e ricos. Barro (2000) confirma que a desigualdade retarda o crescimento em países pobres e aumenta-o em países mais ricos mas, quando se utiliza um conjunto de países alargado, há pouca relação entre desigualdade de rendimento e taxas de crescimento. No entanto, muitos autores assumem que há um impacto da tecnologia sobre a desigualdade. Na verdade, quando alguma empresa adopta uma nova tecnologia, mudará sempre algo no processo de produção. Pode gerar novos tipos de processos mas vai precisar sempre de alguém com as habilitações certas para liderar este novo processo. Esta hipótese é confirmada por Acemoglu (2002a): os recentes desenvolvimentos tecnológicos afectam a organização do mercado de trabalho que contribui para um grande efeito sobre a estrutura salarial. Finalmente, os membros de segmentos com menos formação académica, por terem mais retornos dos seus investimentos em educação, acham vantajoso investir mais em educação, o que contribuirá para um maior crescimento económico. Estes quadros teóricos anteriores são empiricamente confirmados por Barro (2000) e Jaumotte *et al.* (2013). Barro (2000) apresenta estimativas de efeitos fixos e equações do Índice de Gini em variáveis como o PIB e o PIB ao quadrado, escolaridade, índice de democracia, transparência, estado de índice de direito e várias dummies. Nas suas estimativas de efeitos fixos, as dummies para o rendimento ou gastos e ensino secundário, são negativamente relacionadas com a desigualdade e a alta escolaridade e nível de abertura está positivamente relacionado com a desigualdade (com coeficientes significativos). A escola primária e a dummy para os dados individuais ou agregados não têm relação estatisticamente significativa com o coeficiente de Gini. Há uma forte relação de U invertido com o PIB (a chamada curva de Kuznets) nas estimativas de Barro. *Diante dos poucos artigos que estudam os efeitos do capital humano e da tecnologia na desigualdade, em contraste com um debate teórico já bem estabelecido, vamos explorar essas relações nos dois últimos ensaios da tese. As questões da heterogeneidade dos países e da dependência entre eles estão completamente ausentes da literatura sobre os efeitos do capital humano e da tecnologia na desigualdade. O quarto ensaio estima coeficientes heterogéneos de dados em painel para uma especificação da desigualdade no capital humano, tecnologia e grau de abertura e encontra uma forte causalidade e efeito positivo do capital humano na desigualdade, ou seja, quando o capital humano aumenta, a desigualdade aumenta, como a teoria também prevê. O quinto ensaio relaciona várias tecnologias com a desigualdade e encontra fortes indícios de que as tecnologias da informação e comunicação (TIC) mais antigas e as tecnologias dos transportes (e menos frequentemente as tecnologias de informação e comunicação moderna) são habilidades complementares no aumento da desigualdade.*

Abstract

This thesis investigates the relationship between human capital and technology, by one side and the ancestral genetic diversity of populations, by the other. Then, it seeks to understand the determinants of several inventions that occurred during the process of industrial revolutions. Finally, it presents evidence on the determinants of inequality, a clear overlooked issue in the empirical literature. The first chapter, Introduction, provides a critical revision of the literature and stresses the contributions of the thesis. The second chapter highlights theoretically a new channel through which genetic diversity can affect development: human capital. It also shows strong empirical evidence of a hump-shaped relationship between genetic diversity and human capital. This means that some of the human capital achievements today may stem from the genetic diversity mostly determined many centuries ago. Results are robust to the introduction of several controls and to IV estimation. The third chapter investigates the relationship between technology in 1500 and the ancestral genetic diversity of populations. It highlights a strong hump-shaped relationship between genetic diversity and technological developments in 1500. This means that some of the technological achievements may stem from the genetic diversity mostly determined more than a millennium ago. Results are robust to the introduction of several controls, and to IV estimation. The fourth chapter contributes to answer to the question "why inventions that shaped industrial revolutions have been discovered in some countries and not in others?". We access the determinants of more than a hundred inventions around the world, explaining why they occurred in a given country and why some occurred earlier than others. We confirm the importance of scale and dismiss the importance of education as triggers of inventions. Geographic and genetic distance from the UK and the USA have proven to be significant. Both distance from the UK and proximity to the USA seem to have significant effect on the rise of the probability to invent. A fruitful recent theoretical literature has related human capital and technological development with income (and wages) inequality. However, empirical assessments on the relationship are still scarce. In the fifth chapter, we relate human capital, total factor productivity (TFP), and openness with inequality and discover that, when countries are assumed as heterogeneous and dependent cross-sections, human capital is the most robust determinant of inequality, contributing to increase inequality, as predicted by theory. TFP and Openness revealed to be non-significantly related to inequality. These new empirical results open prospects for theoretical research on the country-specific features conditioning the causal relationship from human capital, technology and trade to inequality. In the last chapter of the thesis, we relate technological adoption (of different technologies) with income inequality. We discovered that some technologies such as aviation, cell phones, electric production, internet, telephone, and TV are skill-complementary in raising inequality. We constructed standardized indexes of skill-complementary technological adoption for modern Information and Communication Technologies (ICT), older ICT, production and transport technologies. We found strong evidence that older ICT and transport technologies (and less frequently modern ICT) tend to increase inequality. Additionally, we discovered that results are much stronger in rich countries than in poor ones. Our results are quite robust to a series of changes in specifications, estimators, samples, and measurement of technology adoption. These results may bring insights to the design of incentive-schemes for technology adoption.

Keywords

Genetic Diversity; Determinants of Development; Determinants of Human Capital; Technology; Inventions; Industrial Revolutions; Determinants of Technological Development; Income Inequality; Technological Adoption; TFP.

JEL Codes: I24, I25, I32, N10, N30, O10, O33, O50, Z10.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This thesis contains five essays on economic development, human capital, technology and inequality. The thesis begins with this enlarged introduction which aim is to insert the contribution of the thesis on the literature. Then it proceeds with each essay.[1]

One of the main issues for economists is development. It generates a lot of concerns because it can change the quality of life by creating wealth, employment and minimize social problems. On the other hand unsustainable growth creates income inequality, lack of education, social and environmental problems, etc. The economists seek for the determinants of economic growth and development is as older as sustainable economic growth and rooted to the Industrial Revolution and so to the beginning of the economists' profession. However it is after the mid of the twentieth century that the systematic study of economic growth has begun. Robert Solow has a remarkable role in this process as the author of the most complete exogenous growth model, with scientific repercussions still up-to-date. In fact, the Solow model and its descendants allow us to explain quite well the differences in income *per capita* among different countries. Growth accounting, an empirical technique that allows economists to decompose growth into the contributions of different production factors (such as technology, physical capital and human capital) have also emerged from the Solow model. Barro (1991) systematized a way of evaluating empirically the determinants of economic growth, with an empirical technique known as *growth regressions*. Mankiw *et al.* (1992), known as taking Robert Solow seriously, is another good example of testing his model empirically. These "growth regressions"have evolved along with econometric techniques, addressing problems linked with endogeneity, essentially caused by reverse causation, but also by omitted variables and measurement errors. Simultaneously, the huge increase in data availability allowed enlarging the empirical techniques to panel data analysis and time series analysis. However, causality has been always a major issue in the empirics of economic growth, since most determinants of growth (such as savings, education, and current institutions) may also be a consequence of growth. Endogenous growth theory surged in the late 80s of the twentieth century (first contributions are due to Romer, 1990, Lucas, 1988, Aghion and Howitt, 1992 and Rebelo, 1991) also contributed to the identification of endogenous regressors. In fact, one of the paths economic science used to escape from the causality trap was to identify empirical models with theoretical models.

A very recent literature is centered on the explanation of growth due to institutions (this granted the Nobel Prize to Douglas North in 1993 "for having renewed research in economic history by applying economic theory and quantitative methods in order to explain economic and institutional change"). This has opened a fruitful avenue of research, where the quality of institutions has been associated with economic growth performances of countries and levels of development. Among that literature some recent contributions have also focused on the historically rooted factors that shaped institutions nowadays and constituted the deep determinants of growth and development. Determinants as culture, geography, language, religion, ethnic and genetic composition of the populations, colonial power of a given country have been studied

[1]Throughout the thesis the names essay, chapter, article and paper will be used to name the different chapters of the thesis interchangeably.

as well as their links to current institutions, quality of governments and trade. *The first three essays of this dissertation are inserted in this very recent literature.*

The *deep determinants* of economic growth and development are part of the literature that was born with the Unified Growth Theory (Galor, 2005). This was the first theoretical attempt to encompass an endogenous link between a Malthusian epoch, in which the growth of income and of population are closely tied, and a neoclassical epoch in which growth in physical capital and eventually in human capital and technology yield positive growth rates of income *per capita*. Empirical studies then evaluated the deeply-rooted determinants of different development paths, from which *culture*, religion, institutions, genetic and biogeographic features were evaluated. Guiso *et al.* (2009) tried to capture the effects of culture on economic exchange. They found out that historical and cultural variables, as religious similarity, trust, linguistic common roots or having the same legal origin, affect relations of trust between two countries and also differences in trust tend to positively affect trade, investments and foreign direct investment. On a different point of view, Lada (2013) studied the effects of cultural similarity (using some different measures as racial proximity, religion and civilization) and its effect on conflicts between 1816 and 2009, and found that cultural similarity causes more hostility and increases war-proneness so this generates uncertainty which is adverse to growth. One of the most important studies on this field is Ashraf and Galor (2011a) which tried to obtain historical relationships of cultural assimilation and cultural diffusion (measured by cultural diversity and isolation index) and their relationship to the different patterns of economic development world-wide. Empirically, with OLS (Ordinary Least Squares) and IV (Instrumental Variables) methods using a sample of 60 countries, they concluded that the geographical isolation on pre-industrial era, had a negative impact on cultural diversity on modern era, measured by diversity index from the World Value Survey (WVS) data, but this effect of isolation on cultural diversity has suffered a decrease over the 20th century. Their results indicate that one standard deviation increase of geographic isolation actually decreases cultural diversity by 0.48 standard deviations. Geographical isolation had a positive impact on economic development in the agricultural era, in particular there is a positive effect of the isolation index on the log of population density in years 1 CE, 1000 CE and 1500 CE. Furthermore, geographic isolation had a negative impact on income *per capita* in the industrialization era; specifically the isolation index had a statistically significant and robust effect on the log of income *per capita* in 1960 when most of nations were in the process of industrialization. Cultural diversity had a positive impact on economic development in the process of industrialization, precisely a positive effect of the diversity index on the log of income *per capita* in 1960. Also, *religion* can be an important factor for growth and development since it made civilizations to expand across globe and endure historically and also generated conflicts, in particular between Muslim and Christian societies, which lead to some extent of religious fractionalization within national borders. According to Iyigun (2011) this lack of cooperation is really adverse for economic development. La Porta *et al.* (1999), after their investigation for determinants of the quality of governments across 152 countries using the OLS method, found out that countries which population are predominantly protestant have better governments than catholic or muslin. Furthermore, historically, religion can influence human behavior nowadays, as Andersen *et al.* (2011) concluded, after their empirical study for England between 1377 and 1801. Regions of England that had more Cistercian influence tend to have populations with greater capability of work harder and thrift today because it encouraged a deep cultural change. Barro and McCleary (2003) show that church attendance tend to reduce economic growth because higher attendance means higher allocation of resources (time) to the religion sector, but beliefs in hell, heaven, and after-life tend to increase economic growth, which is explained by the motivation of the population to preserve aspects of behavior that enhance productivity. Theoretical studies were really important for establishing some thoughts about *institutions*. Tabellini (2008) adds that characteristics of political institutions from the past have left a mark on present attitudes and values, e.g., regions where anarchic governments explored their citizens developed a culture of mistrust. Sokoloff and Engerman (2000) tried to find evidence about the differences of development between United States and Canada and the other American countries. Their conclusions suggest that early differences in the degree of inequality in wealth, human capital, and political power look very important because these characteristics may have been preserved by the economic institutions and, somehow, affected growth. So when there was great inequality, privileged elites and weak access of the population to economic opportunities, these members of elites maintain their status over time not generating the full potential of the economy. Following Acemoglu *et al.* (2005) persistence and change in institutions are equilibrium outcomes and both have to be studied as part of the same equilibrium framework. Additionally, policy interventions have an important role in changing institutional equilibrium. Some empirical studies had their relevance on this issue. Acemoglu *et al.* (2001) tried to proof if differences in colonial experience lead to differences in institutions using two different samples (whole world sample and base sample with 64 countries) and two different regressions (OLS and Two-Stage Least Squares - 2SLS). They affirmed that there is a correlation between European settlements and early institutions and also correlation between early institutions and institutions today, so settlements build their institutions and maintained it over time, keeping some of their characteristics nowadays. Later on, Rodrik *et al.* (2004) confirmed these results using part of their sample and established the primacy of institutions over geography in determining the current levels of development. Studying 15 European countries, Agostino and Scarlato (2012) expressed that improvements in political institutions and quality of governance affects positively growth of technology and investment. For environmentally sustainable institutions, there is a positive effect on technology and investment growth, which increases GDP growth. This effect is more significant on countries with high levels of trust and competition, maybe because the population feels more secured to invest when they feel commitment, impartiality and transparency by their government. Dalgaard and Olsson (2013) uncovered a positive link between more politically cohesive countries and income. Olsson (2007) found out that in countries that present weak institutions but abundance in diamonds, a negative relationship between economic growth and the relative abundance of the referred natural resource emerges. In countries with weak institutions there is always the payment of the diamond rent. Some empirical results are contested by Glaeser *et al.* (2004). According to these authors, measurements of institutions have conceptual flaws and is suggested that researchers need to focus on actual laws, rules, and compliance procedures. Specifically what they point out is that indices of institutional quality are highly correlated with each other and with *per capita* income so if institutions improve, as a country grows, it is clear that institutions quality will be higher in richer countries. Another point is that constitutional measures of constraints are correlated with *per capita* income. Finally, the measures of judicial independence and constitutional review are uncorrelated with *per capita* income meaning that these measures are noisy. Hall and Jones (1999) tried to understand why some countries invest more on physical and human capital than others and why some are more productive than others. Their study included 127 countries and their main results show that large differences in income across countries are caused by differences in social infrastructure a measure related to property rights, institutions, and openness to trade which imply large differences in capital accumulation, educational attainment and productivity. A great contribution on this matter, because of the multiplicity of variables and larger samples used, is that from Alesina *et al.* (2003), who concluded that ethnic and linguistic fractionalization, which can determine the social cohesion and the social interactions, are essential determinants of economic growth but it is difficult to evaluate the size of this relationship because of the strong correlation with other variables. Lindner and Strulik (2008) affirmed that social fractionalization (different groups in the society with different languages, beliefs, ethnic origin, etc.) presents a nonlinear correlation with risk of expropriation and has a negative impact on economic growth. However, social fractionalization has a negative impact on growth only if secured property rights are not enforced. Insecure property rights and slow growth are more present on countries with sufficiently polarized societies which suffer from low social capabilities. Ahlerup *et al.* (2009) tested a theoretical model, by means of cross-country growth regressions, in which they found that the marginal effect of social capital (interpersonal trust) decreases with the strength of the institutions of the country. *Population diversity* can be helpful for productivity because it is quite connected to the theory of skills complementarity. In fact, different workers have different ways to react to certain problems at work, which affects their productive skills. In fact, diversity may cause divergent opinions, beliefs and preferences which may generate lack of trust, communication barriers or even conflicts. Ashraf and Galor (2013), with the objective of finding evidence of *genetic diversity* as important determinant of economic development, is one of the most important works on this field. They used population density and income *per capita* as dependent variables, genetic diversity as main independent variable and many controls and instrumental variables which made their results more robust. They verified that the exodus of *Homo sapiens* out of Africa affected genetic diversity and had historically long lasting effect on the pattern of comparative economic development. Genetic diversity has a hump-shaped effect on economic development in the precolonial era. Genetic diversity has a non-monotonic effect on income *per capita* in the modern world. Low and high level of diversity are disadvantageous for development but, on the other hand, intermediate levels of diversity cause more development. Seeking for effects of geographic, historical and cultural variables on income and productivity, Spolaore and Wacziarg (2013a) stated that genetic distance and its effects on income differences reached the peak in the second half of the nineteenth century and started to decline at the second half of the twentieth century. They also report an indirect and persistent time effect of pre-historical biogeographic factors which is also documented by Ashraf and Galor (2011b). Ethnic fractionalization appears to have some influence on economics, following the thoughts of Rohner *et al.* (2013a). Population who lives at places with more violence tend to increase ethnic identity and conflicts seem to strengthen solidarity within the ethnic group. Intensity of fighting has a negative effect on the economic situation on highly fractionalized countries, because the population which is fighting is not cooperating and stimulating the economy, but on less fractionalized countries it has no effect, result also reached by Rohner *et al.* (2013b). The reverse effect is also true, Caselli and Coleman (2013) found evidence that economic development can encourage the decrease of ethnic conflicts. For Sub-Saharan Africa analysis and for three different decades (60s, 70s, and 80s), Easterly and Levine (1997) determined that Africa's poor growth and income are associated with political instability, distorted foreign exchange markets, poor schooling, underdeveloped financial systems, and not enough infrastructures. Furthermore low schooling, underdeveloped financial systems, distorted foreign exchange markets, and insufficient infrastructure are closely associated with high ethnic diversity, confirming the models according to which highly fractionalized societies are more prone to competitive rent-seeking, disagreeing most of the times, on public goods or infrastructures and education. Taking into account a more diverse study, which includes 195 different countries between 1990 and 2000, and using birthplace diversity, a variable never used before, Alesina *et al.* (2012) found out a positive effect of birthplace diversity of nationals and of immigrants on productivity, relation which is stronger in highly productive countries. Moreover there is also evidence of positive effects of birthplace diversity for unskilled workers on productivity. Population diversity can generate good outcome but can create unpleasant situations as Spolaore and Wacziarg (2013b) realized. Populations that are genetically closer tend to have more conflicts even after controlling for geographic distance, income differences and other variables related to conflict. Populations which share recent common history tend to converge more in preferences and characteristics and also are more prone to fight. On the other hand, Easterly and Levine (2012), focusing on the impact of European share of population during the early stages of colonization on economic development today, concluded that Europeans influence had a lasting and positive effect on economic development. These findings hold when: the sample is restricted to non-settler colonies, proportion of the population of European descent is conditioned, and euro share is instrumented. *Biogeographical* and geographical factors can condition, positively or negatively the development of a country. Gallup and Sachs (2001), trough cross-country regressions between 1965 and 1990, have found a negative relationship between malaria and economic growth, even when controlling for many factors, like location, history, other tropical diseases, and isolation. Olsson and Hibbs Jr. (2005) modelled the transition from a hunter-gatherer economy to an economy based on agricultural, claiming this transition to be a crucial event that turn possible the Industrial Revolution, since in the first place led to technological progress. The authors also stress the importance of initial geographical and biogeographical factors to the location and timing of these transitions. They test these hypotheses empirically and found significant results for many countries, confirming the existence of a relationship between biogeographical and geographical factors and development. For Comin *et al.* (2013) spatial distance is a determinant for diffusion, countries as that are far away from adoption leaders have a slowly adoption but the spatial effects vanish over time.

One of the conclusions of this critical review is that most of the deeply rooted determinants of development literature have analyzed empirically the effect of such determinants on proxies of development such as population density or per capita income. Only a minority have addressed the effect of those historically rooted determinants on human capital (e.g. Hall and Jones, 1999) or technological progress (e.g. Comin et. al. 2013). Our first three essays are contributions to fill in this gap. In particular, we study the effect of genetic diversity on several measures of human capital (1st essay), on several measures of technological adoption (2nd essay) and analyze the determinants of technological inventions during the process of industrial revolution (3rd essay). This thesis highlights a robust hump-shaped relationship between genetic diversity and the current levels of human capital and between genetic diversity and the pre-colonial level of technological adoption. Moreover, it also confirms the importance of scale effects and distance to USA and UK as important triggers of innovations during the process of industrial revolution.

Furthermore, the origin of income inequality seems to be a great concern to economists because, besides of being hard to find its roots, it is sensitive to some changes which are generally accepted and encouraged as the education or growth. The theoretical approach for inequality and level of *per capita* income contains some version of the Kuznets (1955) curve, which developed a model that builds on the reallocation of persons and resources from agriculture to industry. Later on, this relationship started to be described as an inverted-U; i.e, inequality rises until it reaches some point from which it falls again, as an economy develops.[2] Allen (2003) tested six major explanations for the reverse in the progress of the Mediterranean versus the United Kingdom and the Netherlands in early modern Europe - population, enclosure of the land, the role of the empire, the role of the government, technology, and literacy; and found the role of empire (via international trade) to be the most important factor to the reverse of development and the increase of inequality in the Mediterranean countries. Allen (2009b), with a literature review about Industrial Revolution and its effect on inequality, realized that in the first four decades of the 19th century, inequality increased considerably, the profit rate doubled and the average wage stagnated and finally growth cannot be explained without inequality because they influenced each other. This relationship seems to be stronger when countries are divided into two groups: poor and rich countries. Barro (2000) confirms that inequality delays growth in poor countries and encourages it on richer countries but, when using broad panel of countries, there is little relation between income inequality and rates of growth. In fact, when restricted to groups or to one single country the relationship between growth and inequality seems to generate more comprehensive results. A study by Birchenall (2001), which analysed the increase in income inequality in Colombia and related it to growth, concluded that the Colombian economy had slow growth when it has unequal and immobile conditions. Chakrabarti (2000) proves, in an attempt to examine the effect of international trade on income inequality, and for 73 different countries, that growth establishes a channel through which trade decreases inequality. Neves and Silva (2014) surveyed the empirical literature on the effects of inequality on economic growth. The authors conclude that the effects are dependent on countries, variables used to measure inequality, time periods, structure of data, and econometric techniques. Until now, we reviewed evidence of the effect of inequality on economic growth. However, many authors assume that there is an impact of technology on inequality. In fact, when some company adopts a new technology it will always change something on the production process. It can generate new types of processes, but it will always need someone with the right skills to lead this new process. This hypothesis is confirmed by Acemoglu (2002a): recent technological developments affected the organization of labour market which contributed for a large effect on the wage structure. Theoretically speaking, Acemoglu (2003) assumed that as long as market size for skill-complementary technologies expanded, the generation of new and more technologies became more profitable, explaining the increase in the demand for skills and, consequently, the rise in the returns to education and inequality. This statement is confirmed by Galor and Moav (2000), which concluded that an increase in the rate of technological progress increases return to ability and also raises the wage inequality between skilled and unskilled workers. Additionally, Aghion *et al.* (2002) claim that an increase in technology increases long-run within group inequality which exceeds its new steady-state level. This effect reacts positively to an increase in most of technologies and to physical capital. Adding two externalities to their theoretical model, home environment externality (parental education importance for formation of human capital of the child) and global technological externalities (ability of individuals to adapt to new economic environment or technology change), Galor and Tsiddon (1997) affirmed that those lead to the progress of income distribution, human capital distribution wage differential, and economic growth. This happens because technological progress affects more the productivity of high-skill individuals than of low-skill individuals but this also increases wage inequality. Finally, members of the less-educated segments, because

[2]Gutiérrez and Tanaka (2009), for example, found this relation on education and inequality. They concluded that there is a threshold level of inequality above which public education suffers a reduced support, leading to less expenditure per student and low tax rates.

of having higher returns from their education investment, find advantageous to invest more on education, which will contribute for more economic growth. These theoretical frameworks above are empirically confirmed by Barro (2000) and Jaumotte *et al.* (2013). Barro (2000) presents fixed-effects estimations of equations of the Gini index on covariates such as GDP and GDP squared, schooling, democracy index, openness, rule of law index and several dummies. In his fixed-effects estimations, dummies for income or spending and secondary schooling are negatively related to inequality and higher schooling and openness are positively related to inequality (with significant coefficients). Primary schooling and the dummy for individual or household data are insignificantly related to the Gini coefficient. There is a strong inverted-U relationship with GDP (the so-called Kuznets curve) in Barro's estimations. Jaumotte *et al.* (2013) studied 51 advanced economies from 1981 to 2003, concluded that technological progress tend to increase the demand for skills and education, which contributes for demand of higher salaries and inequality in income, in particular FDI tends to take place in higher technology sectors, resulting in the increase of employment and income but only to those that already have higher skills and education promoting inequality. A similar finding was obtained by Caselli (1999), which concluded that the increase in wage inequality is related with the increase in inequality in capital-labour ratios, since industries with higher rates of adoption will attract more highly paid workers and industries with low adoption rates will have workers with lower wages. Finally, Ding *et al.* (2011), which analysed the impact of improved upland rice technologies, promoted by the government in rural China, on income inequality, established that households which adopted the improved upland rice technology, had more income, approximately 14 to 16% more, than non-adopters. These results are contrary to the findings of Lin (1999) which studied 500 households from five counties in China. His major finding is that when a new rice technology is available, the households which are adopting it change their focus to production of rice, leaving all the other production types they might have.

In face of the very few articles studying the effects of human capital and technology on inequality, in broad samples of countries an issue with an established theoretical debate, we explore those relationships in the two last essays of the thesis. Issues of country heterogeneity and dependence are completely absent from the literature on the effects of human capital and technology in inequality. The 4^{th} essay estimate heterogeneous panel data estimators for a specification of inequality on human capital, technology and openness and conclude for a strong causal and positive effect of human capital in inequality, i.e. as human capital rises, inequality increases, as theory also predict. The 5^{th} essay relates several technologies with inequality and found strong evidence that older information and communication technology (ICT) and transport technologies (and less frequently modern ICT) are skill-complementary in increasing inequality.

Chapter 2

Human Capital and Genetic Diversity

2.1 Introduction

The determinants of human capital have been reported very little in the literature. Although there is a huge literature on the determinants of schooling linked with quality of schooling, finding a weak causality between inputs to schooling and educational achievements, there are not many contributions that explore the deeply-rooted determinants of investment in the quantity and quality of human capital. In fact, Hanushek and Woessmann (2011) concluded for very little evidence for the influence of inputs in cross-country differences in schooling achievements while recognizing the essential effects of family and social background and of some country-specific school-related institutions.[1] However, the literature on the deeply-rooted determinants of development has been greatly developed in the last decade. Hall and Jones (1999) demonstrated the importance of social infrastructure, a composed measure of law of rule and other institutional measures, which was then followed by Glaeser *et al.* (2004). It is interesting to note that Glaeser's paper highlights the role of human capital as a more significant source of growth than institutions but does not show any causal relationship between institutions and human capital. Rodrick *et al.* (2004) showed evidence for the supremacy of institutions related to geography and integration as determinants of economic development. However, Olsson and Hibbs Jr. (2005) showed the strong influence of geographical and biogeographical factors in determining the current level of development. In related literature on institutions, Sokoloff and Engerman (2000) and Acemoglu *et al.* (2005) have stressed the role of colonialism, while the effects of ethnolinguistic fractionalization were examined by Easterly and Levine (1997) and Alesina *et al.* (2003).[2] Moreover, the historical impact of sociocultural factors has been highlighted by Barro and McCleary (2003), Tabellini (2008), and Guiso *et al.* (2009).

Ashraf and Galor's (2013) work is probably one of the most influential recent papers in the field of economic development.[3] Their paper reports a significant relationship between genetic diversity determined ancestrally and economic development in the present. The paper illustrates the relationship between genetic diversity and development through the positive and decreasing effect of genetic diversity on technology and a negative effect of genetic diversity on output representing inefficiency in production. The hump-shaped empirical relationship exposed by the authors builds on two opposite effects of genetic diversity. First, an increase in diversity enhances production possibilities as a wider spectrum of traits is more likely to contain those that are complementary to the advancement of superior technologies. In fact, some competition for survival, as natural selection explains, also increases adaptability and improves the society's ability to successfully introduce new and better technologies. However,

[1]For example, Sequeira and Ferraz (2009) and Sequeira (2009) highlight the significant effect that country-risk has in education measures.

[2]The influence of genetic diversity on ethnolinguistic fractionalization has been studied by Ahlerup and Olsson (2012).

[3]The paper was a lead article in the American Economic Review and was the Science's Editor Choice in September 2012.

after a certain level of genetic diversity, further rise would increase the scope for disarray and mistrust, increasing the probability of conflict.

The aim of our paper is to purpose an alternative channel through which genetic diversity can influence development: human capital. In fact, wider variety of traits enhance the capacity to accumulate different types of skills; also after a certain level of genetic diversity the cost to accumulate knowledge increases due, e.g., to the adaptation of curricula and school routines to serve special needs and diverse cultural background and the possibility of conflicts between groups in the school context. The relationship between human capital and genetic diversity will emerge from the interplay of those opposing forces. Our paper exhaustively describes the empirical relationship between several measures of human capital and genetic diversity and tests its robustness in a cross-section of countries.

In Section 2.2 we revise the literature that relates genes and particularly genetic diversity with human capital. In Section 2.3 we devise a simple model in which the idea of the relationship between genetic diversity and human capital is presented. In Section 2.4, we present the main empirical results, subject to a number of extensions and robustness tests. Section 2.5 concludes.

2.2 Genetic Diversity and Human Capital: A Literature Review

The influence of genetic traits on school achievement has been discussed in the psychology and medicine sciences. The behavioral sciences have emerged from an era of strict environmental explanations for differences in behavior to a more balanced view that recognizes the importance of nature (genetics) as well as nurture (environment). On the evaluation of the relative strength of nature and nurture in school achievements, Petril et. al. (2004) presented evidence on intraclass correlations which reflected both considerable genetic influence at each age and modest shared environmental influence within and across ages. Plomin and Craig (2001) supports that general cognitive ability, a key factor in learning and memory, is among the most heritable behavioral traits. General cognitive ability is partially determined by a network of genes. Thus, the authors argue that multivariate genetic research contradicts the idea that genes work on specific cognitive processes but are instead complementary in forming the general cognitive ability. Hill et. al. (2014) showed that differences in general cognitive ability (intelligence) account for approximately half of the variation in any large battery of cognitive tests and are predictive of important life events including health. According to the same authors, genome-wide analyzes of common single-nucleotide polymorphisms (SNP) indicate that they jointly tag between a quarter and a half of the variance in intelligence. However, no single polymorphism has been reliably associated with variation in intelligence. It remains possible that these many small effects might be aggregated in networks of functionally linked genes. Rietveld et. al. (2013) showed three (highly-) statistically significant SNPs in explaining differences in schooling. A linear polygenic score from all measured SNPs accounts for near 2% of the variance meaning a modest effect of near 1 month *per* allele. De Neve *et al.* (2013) discovered that the leadership tendency may be associated with a SNP residing on a neuronal acetylcholine receptor gene (CHRNB3). Shin (2014) showed significant differences on scores obtained on the *Korean version of modified Mini-Mental State Examination* between heterozygotes and homozygotes participants (regarding the presence of the apolipoprotein polymorphism), effects that vary across ages and levels of education. As human capital is a complex concept consisting on different skills that shape human abilities, this reinforces the argument towards the comple-

mentarity of genetic traits on a global (and maybe non-linear) influence on human capital.
Although the relative quantitative effect of genetics on intelligence and school achievements will certainly continue on discussion in the genetic-related fields, there is now some evidence of the (positive) influence of a network of genetic traits on defining the human ability to learn. Additionally, genetic influence on achievements may not be independent on social background as Tucker-Drob and Harden (2012) had shown. However, the exploration of non-linear effects of heterozygosity (genetic diversity) on school achievements at the micro-level is clearly overlooked, as the only thinly related study is that from Shin (2014), already mentioned.
Despite the inexistence of direct evidence of a negative effect of genetic diversity in the schooling outcomes, it is not difficult to anticipate such an effect. In the following lines, we outline evidence for the (mostly indirect) effects of genetic diversity in school outcomes. For example, Adobo and Agbayewa (2011) show that homogeneous ability level grouping is superior for promoting students learning outcomes. Carter (2003) describes the effects of "black" cultural capital in minority students outcomes. Thus, higher group diversity has a negative effect in schooling outcomes. May this group diversity include genetic diversity? One may argue that the cultural differences between ethnic diverse groups are more due to nurture than to nature. Some argue however that cultural behavior may also have genetic roots (see e.g. Pyysiäinen and Hausser, 2010 who argue that religion may have genetic roots). Some genetic-rooted learning disorders which imply costs to be overcome have also quite different prevalence rates by different ethnic groups. Examples are autism, attention deficit and hyperactivity syndrome and handedness (see e.g. Koriath, 2014, Akinbami et. al., 2011; Johnston et. al., 2009 and McManus, 2009). Another good example is violence as a consequence of genetic diversity. In fact, Ferguson and Beaver (2009) stressed that violence is a product of evolution and identified a number of polymorphisms associated with violence. Genetic diverse societies are also the adequate environment to trigger the natural selection predisposition to violence. It seems that the way to address the difficulties in learning in diverse genetic environments has been to increase the efforts of educators to compensate the additional costs to education imposed by diversity (see e.g. Terry and Irving, 2010 and Ponciano and Shabazian, 2012).

2.3 Genetic Diversity and Human Capital: A Simple Model

Consider an illustrative model of an economy where the supplied stock of human capital depends from family investment h and from inherited general cognitive ability $F(\omega, x_i)$, a set of complementary genetic traits (where $0 < \omega < 1$ is the degree of genetic diversity) which shape abilities (x_i) to read and write, logical and calculus skills, creativeness, ability to work in group or to lead. General cognitive ability is specified as:

$$F(\omega, x_i) = \left[\int_0^{\omega} x_i^{\theta} \right]^{\alpha}, \text{ with } 0 < \alpha < 1, \text{and } 0 < \theta < 1 \quad (2.1)$$

where $F(\omega, x_i) > 0$ is an increasing concave function ($\frac{\partial F(\omega)}{\partial \omega} > 0$; $\frac{\partial^2 F(\omega)}{\partial \omega^2} < 0$) that guarantees the complementarity between the different traits in shaping the general cognitive ability. The stock of human capital can be written as:

$$H = h \left[\int_0^{\omega} x_i^{\theta} \right]^{\alpha}. \quad (2.2)$$

The family decides consumption, *c*, and human capital investment, *h*, in order to maximize a logarithmic utility as $U(h,c) = log(h) + \beta log(c)$ subject to a resource constraint $p_c c + p(\omega)h = y$. The family cannot influence the genetically-inherited general cognitive ability but decide hers own investment in education given the general cognitive ability F. The price of education $p(\omega) > 0$ is an increasing non-concave function of genetic diversity ($\frac{\partial p(\omega)}{\partial \omega} > 0; \frac{\partial^2 p(\omega)}{\partial \omega^2} \geq 0$). This feature of the model incorporates the idea according to which increasing genetic diversity in school environment implies diverse genetic-rooted behaviors (e.g. religious aspects of the daily life, violent behaviors) and an increased probability of having at least one element with learning disabilities. Those features of increased diverse school environments represent a cost of learning associated to adaptation of curricula and routines to people coming from different origins, special education services target at children with special needs, policing services to deter potential violence in schools. As is natural, the properties of $p(\omega)$ guarantees that these types of learning costs are increasing more than proportionally to the degree of genetic diversity. This means e.g. that the costs of increasing genetic diversity from a relatively homogeneous situation are quite low. Although the family do not control genetic diversity, she becomes aware of the cost of education associated with it. First order conditions, after solving in order to h, yields

$$h = \frac{y}{(1+\beta)p(\omega)}. \quad (2.3)$$

Substituting equation (2.3) into (2.2) yields

$$H = \frac{y}{(1+\beta)p(\omega)} \left[\int_0^\omega x_i^\theta \right]^\alpha. \quad (2.4)$$

Interestingly, this simple model encompasses important features linking human capital and genetics described above. First, it predicts an influence of family income on average human capital, which is highlighted by all the human capital theory. Taken broadly, y can be taken as the family endowments and thus can be interpreted as including also parents education. Second, it includes an inherited composed measure of complementary genetic traits, generally defined as general cognitive ability. Third, it highlights the costs to education of great genetic diversity. Now, we want to show that equation (2.4), representing the average supply of human capital, is such that it has a maximum for a level of genetic diversity $0 < \omega < 1$. The first-order condition for this maximization problem yields $-p(\omega)\frac{\partial p(\omega)}{\partial \omega} + \alpha \frac{x_i}{\left[\int_0^\omega x_i^\theta\right]} = 0$. Without loss of generality, we assume that each trait has equal contribution to the general cognitive ability and thus $x_i = \bar{x}$. Using this generalization the first-order condition is simplified to:

$$-\frac{\partial p(\omega)}{\partial \omega}\frac{\omega}{p(\omega)} + \alpha = 0 \quad (2.5)$$

which imply that $0 < \frac{\partial p(\omega)}{\partial \omega}\frac{\omega}{p(\omega)} < 1$, meaning that the percentage variation of the price due to a percentage variation in genetic diversity must be equal to $0 < \alpha < 1$.
Second-order condition to a maximum is the following:

$$-\left[\left[\frac{p(\omega) - \frac{\partial p(\omega)}{\partial \omega}\omega}{p(\omega)^2}\right]\frac{\partial p(\omega)}{\partial \omega} + \frac{\omega}{p(\omega)}\frac{\partial^2 p(\omega)}{\partial \omega^2}\right], \quad (2.6)$$

which is always negative at the extremum given by (2.5), given the properties of $p(\omega)$.[4]
In the empirical part of the paper we find a robust value of $\omega = 0.7$ as the proportion of genetic diversity that maximizes human capital, meaning that above this value, more genetic diversity will deter human capital. We will give some examples of functional forms for $p(\omega)$ for which a reasonable value of the share of the general cognitive ability α may replicate empirically the proportion of genetic diversity that maximizes human capital.

Example 1. Consider that $p(\omega) = \frac{1}{1-\omega}$. Equation (2.5) yields $\omega = \frac{1}{1+\alpha}$ which is between 0 and 1. With $\alpha = 0.4286$, genetic diversity that maximizes human capital is $\omega = 0.7$.

Example 2. Consider that $p(\omega) = 1 + \omega$. Equation (2.5) yields $\omega = \frac{\alpha}{1-\alpha}$ which yields $\omega = 0.7$ for $\alpha = 0.4117$.

Example 3. Consider that $p(\omega) = e^{\omega}$. Equation (2.5) yields $\omega = \alpha$ which yields $\omega = 0.7$ for $\alpha = 0.7$.

Given the evidence on the weight of nature in explaining ability exposed in the previous section, those values for α seems reasonable.
If the beneficial effects of genetic diversity dominate at lower levels of diversity and the detrimental effects prevail at higher ones (i.e., if there are diminishing marginal returns to both diversity and homogeneity), the theory would predict a hump-shaped effect of genetic diversity on human capital throughout the development process.

2.4 Empirical Findings

2.4.1 Data and Sources

In this section we describe the variables and data sources for this work. Our dependent variable is human capital, for which we use different measures - enrollments, attainments, scores on international tests, and also measures of quality-adjusted stocks of human capital (the product of scores and quantities of human capital). These alternative measures of human capital were taken from Cohen and Soto (2007) for measures of quantity of human capital and from Hanushek and Woessmann (2012) for measures of the quality of human capital. We also use a measure of the relational capacity of human capital, using trust as a weight for human capital. Trust was taken from Ashraf and Galor (2013), which used data from the World Values Survey conducted during the period 1981-2008.
Explanatory variables rely on the database from Ashraf and Galor (2013) who include variables that measure genetic diversity, as adjusted to migratory movements and ancestrally adjusted. For our benchmark analysis we use the ancestrally adjusted (to 2000) predicted genetic diversity. As genetic diversity measured by this variable has been adjusted to account for the 2000 composition of populations (that can trace their ancestral origins to different source countries in the year 1500) this is the appropriate measure to relate to the distribution of human capital

[4]This is a model for the supply of human capital, as empirical evidence bases on variables mostly linked with the supply side of the market for human capital. If one plug y from the model of Ashraf and Galor (2013) in equation (2.4) the same hump-shaped relationship between human capital and genetic diversity would appear given certain conditions. So this model highlights an alternative channel between genetic diversity and development, through human capital. More generally, any model with an ω -increasing concave general cognitive ability $F(\omega, x_i) > 0$ and an ω -increasing non-concave $p(\omega) > 0$ would yield a maximum for ω fulfilling $\frac{\partial F(.)/\omega}{\partial p(.)/\omega} = \frac{F(.)}{p(.)}$.

Table 2.1: Human Capital Variables

Variables for Human Capital (HC)	Name	Measure (years and source)
Years of schooling of population+25	School	1960-2010 (Ashraf and Galor, 2013)
Social Capabilities of HC	kskh	Trust*School (Ashraf and Galor, 2013)
% of population + 15 with secondary education	m_sec15c	1960-2010 (Cohen and Soto, 2007)
Years of schooling of population +15, inc. students	m_tyr15	1960-2010 (Cohen and Soto, 2007)
Years of schooling of population 15-64	m_tyr1564	1960-2010 (Cohen and Soto, 2007)
Average test score in math and science	cognitive	primary through secondary, all years 1964-2003 (Hanushek and Woessmann, 2012)
Average test score in math and science	lows	lower secondary, all years 1964-2003 (Hanushek and Woessmann, 2012)
Share of students reaching basic literacy	basic	average test scores in math and science, primary through secondary, all years 1964-2003) (Hanushek and Woessmann, 2012)
Share of top-performing students	top	based on average test scores in math and science, primary to secondary school, all years 1964-2003 (Hanushek and Woessmann, 2012)
Interaction between School and Cognitive	Schoolxcognitive	–
Interaction between School and lows	Schoolxlows	–
Interaction between School and basic	Schoolxbasic	–
Interaction between School and top	Schoolxtop	–

in the world after the year 2000. Other explanatory variables will be introduced as controls later on. The countries included in our cross-section analysis depend on the availability of human capital data and human capital variables.[5]

Table 2.1 summarizes the dependent variables used, which seek to measure quantity and quality of human capital, as well as quality-adjusted measures of years of schooling. The variables *school*, *m_sec15c*, *m_tyr15* and *m_tyr1565* measure quantity of schooling (years of schooling or attainment). Variable *kskh* measures schooling adjusted for social capital. We use this variable as a proxy for the relational capability of the existing human capital. The next four variables (*cognit*, *lows*, *basic*, *top*) measure quality of schooling, available as tests scores and share of students reaching certain levels of quality on international tests. The final four variables measure human capital (quantity) weighted by quality (scores). We have also tested other quality-weighted human capital variables, in which we substituted *School* by *m_sec15c*, *m_tyr15* and *m_tyr1565*, alternatively. As results are quite similar to those obtained when using the *School* variable, we choose not to report them. These results are available upon request. Table 2.2 reports descriptive statistics for the dependent variables. The explanatory variables used - predicted genetic diversity ancestry adjusted and mobility index-predicted genetic diversity ancestry adjusted, are *pdiv_aa* e *pdivhmi_aa*, respectively and measure, as explained above, genetic diversity for 2000.

2.4.2 Results: Main Specification

In this section we document the fact that there is a hump-shaped relationship between different variables linked with human capital (quantity, quality, and quality-adjusted measures) and the genetic diversity of countries. We do this estimating OLS regressions.[6] We present figures (Figures 2.1 and 2.2) and a table (Table 2.3), revealing a non-linear relationship (in a hump-shaped or inverted-U form) between human capital variables and genetic diversity. The significance of the squared term of genetic diversity to the adjustment is reinforced with the

[5]The largest list of countries used (123) is detailed in the Appendix. Detailed lists of countries for each regression are available upon request.

[6]In the robustness Section we disregard the potential endogeneity problem, especially related with reverse causality.

Table 2.2: Descriptive Statistics

Variable	Mean	Std. Dev.	Min	Max
(a) pdiv_aa	.7267	.0269	.6279	.7743
(b) pdivhmi_aa	.7229	.02904	.6178	.7826
(1) school	4.8623	2.8126	.4089	10.8622
(2) kskh	1.8560	1.4717	.0661	5.7926
(3) m_sec15c	.1528	.1303	.0088	.5103
(4) m_tyr15	5.6471	3.0635	.5183	11.7767
(5) m_tyr1564	5.8586	3.2381	.5417	12.2517
(6) cognitive	4.5429	.5709	3.0893	5.3376
(7) lowsec	4.5369	.6120	2.6830	5.5116
(8) basic	.7569	.2045	.1817	.9738
(9) top	.0592	.05268	.0001	.2043
(10) school**x**cognitive	30.6408	13.7546	7.8473	53.2535
(11) school**x**lowsec	30.6199	13.8527	7.9422	53.5439
(12) school**x**basic	5.3417	2.7852	.9131	9.9741
(13) school**x**top	.46392	.4049	.0013	1.3676

Notes: (a) Predicted genetic diversity ancestry adjusted; (b) Mobility index-predicted genetic diversity ancestry adjusted. Ancestry adjustment is made to make variables consistent to time-measurement in 2000. Details are given in the Appendix F of Ashraf and Galor (2013).; (1) Years of schooling; (2) Years of schooling∗Interpersonal Trust; (3) % of population aged 15 or over with complete secondary education; (4) Years of schooling of population 15 and over, whether studying or not; (5) Years of schooling of population 15-64 who are not studying; (6) Average test score in math and science, primary through end of secondary school, all years (scaled to PISA scale divided by 100); (7) Average test score in math and science, only lower secondary, all years (scaled to PISA scale divided by 100); (8) Share of students reaching basic literacy (based on average test scores in math and science, primary through end of secondary school, all years); (9) Share of top-performing students (based on average test scores in math and science, primary through end of secondary school, all years); (10) Years of schooling∗Average test score in math and science, primary through end of secondary school; (11) Years of schooling∗Average test score in math and science, only lower secondary; (12) Years of schooling∗Share of students reaching basic literacy (based on average test scores in math and science, primary through end of secondary school; (13) Years of schooling∗Share of top-performing students (based on average test scores in math and science, primary through end of secondary school, all years).

high significance of an F-test to the coefficient of squared genetic diversity, also presented in Table 2.3. This is common to the different measures of human capital that we used. Moreover and interestingly, this rough regression through different dependent variables predicts surprisingly similar maximum values for the genetic diversity above which human capital tends to decrease (around 0.70), which is slightly below the median value for these variables (see Table 2.2). This means that there is a quite realistic value for genetic diversity below which human capital increases with diversity and above which it decreases. Moreover this result is very similar to that obtained by Ashraf and Galor (2013: Table 1) when testing the relationship between genetic diversity and development (using the log of population density as dependent variable). According to our estimates a 1 percentage point increase in genetic diversity for the country with least genetic diversity in the sample would imply an increase in schooling of around 1.48 years (or 0.75 points in cognitive score) and the same increase for the country with highest genetic diversity implies a reduction of roughly 1.42 years of school (or 0.60 points in cognitive score). These are sizable quantitative effects, representing almost 1/3 of the schooling average value and nearly 15% of the tests scores' average. Moreover, the positive effect of increased genetic diversity for lower levels is relatively higher than the negative effect of increased genetic diversity for higher levels. The inclusion of continent dummies would cause in our case a reduction of the positive effect of genetic diversity.[7] If this would be the case, a 1 percentage point increase for the least diverse would cause an increase of nearly 0.56 years in year of school (or 0.31 in cognitive score) and would cause a decrease of almost 0.19 year of school (or 0.26 in cognitive score) for the most diverse. We will further discuss the effect of introducing continent dummies in the robustness section.

[7] As in Ashraf and Galor (2013), the use of continent dummies intends to test against omitted continent-specific features (such as broad cultural heritage or climate).

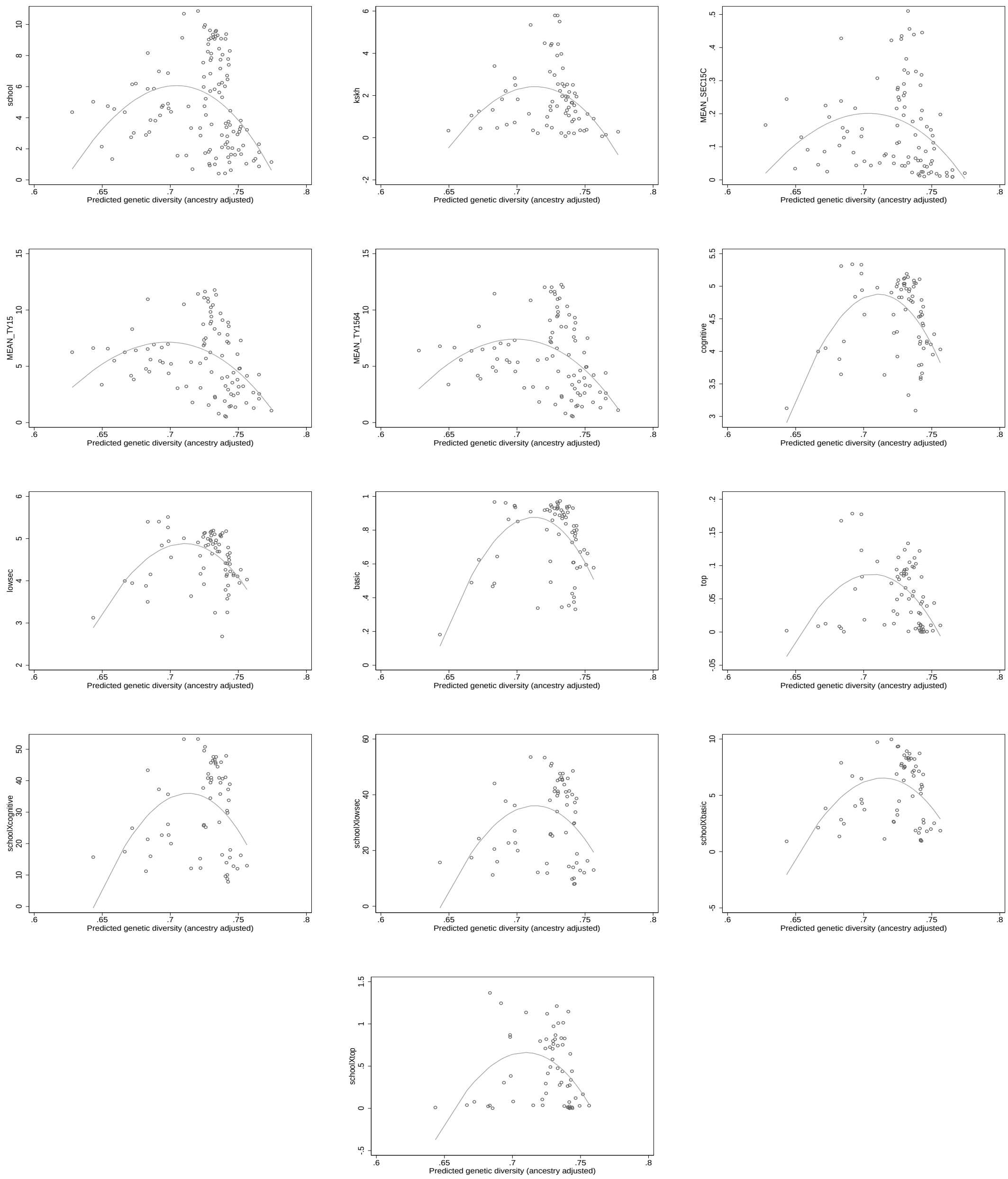

Figure 2.1: The Hump-Shaped Relationship between Human Capital and Genetic Diversity

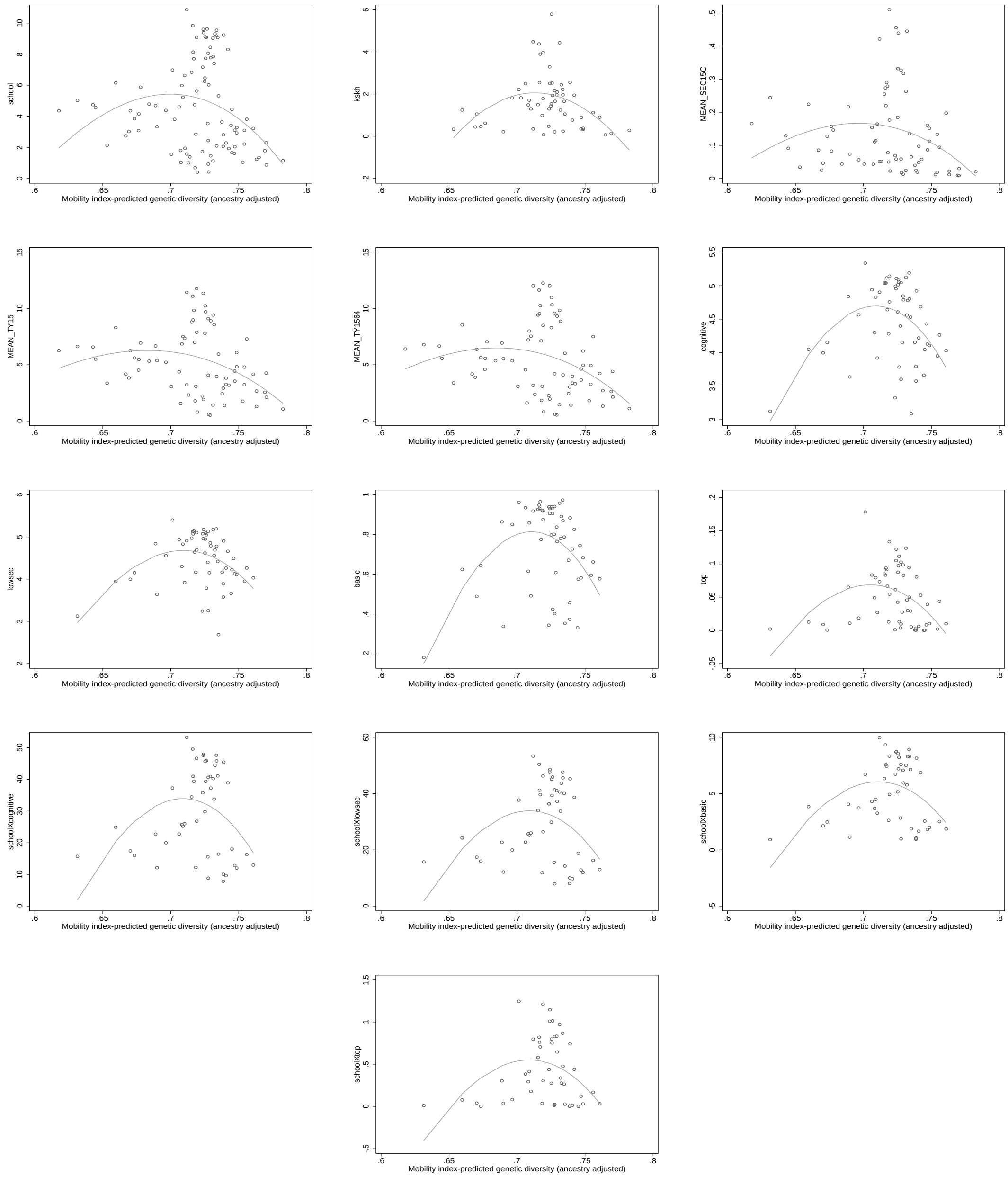

Figure 2.2: The Hump-Shaped Relationship between Human Capital and Genetic Diversity (Mobility Adjusted)

Table 2.3: Restricted Regressions

Dep. Var.	Diversity Var. (x)	Coef. for x ($s.e$)	Coef. for x^2 ($s.e$)	F-test for x^2 (p-val)	$Adj.$ R^2	N
(1) school	pdiv_aa	1391*** (329,7)	-989,5*** (231,0)	18.34 (0.000)	0.11	123
(2) kskh	pdiv_aa	1129*** (236,3)	-793,0*** (165,7)	22.92 (0.000)	0.15	73
(3) m_sec15c	pdiv_aa	47,9*** (15,4)	-34,2*** (10,8)	10.02 (0.002)	0.07	94
(4) m_tyr15	pdiv_aa	1265*** (333,0)	-909,2*** (233,7)	15.13 (0.000)	0.13	94
(5) m_tyr1564	pdiv_aa	1351*** (356,6)	-969,4*** (250,2)	15.01 (0.000)	0.13	94
(6) cognitive	pdiv_aa	654*** (90,7)	-460,9*** (64,3)	51.32 (0.000)	0.25	71
(7) lowsec	pdiv_aa	668*** (96,7)	-470,7*** (68,6)	47.02 (0.000)	0.22	71
(8) basic	pdiv_aa	243*** (33,4)	-170,9*** (23,7)	51.91 (0.000)	0.27	71
(9) top	pdiv_aa	47,5*** (10,6)	-33,7*** (7,5)	20.26 (0.000)	0.20	71
(10) school**x**cogn.	pdiv_aa	10806*** (3883)	-7585,1*** (2736)	7.69 (0.007)	0.12	63
(11) school**x**lowsec	pdiv_aa	10939*** (3910)	-7680,8*** (2754)	7.78 (0.007)	0.12	63
(12) school**x**basic	pdiv_aa	2509*** (734)	-1759,0*** (517,1)	11.57 (0.001)	0.17	63
(13) school**x**top	pdiv_aa	350,5*** (97,9)	-247,4*** (69,1)	12.84 (0.001)	0.13	63
(1) school	pdivhmi_aa	790,1*** (216,4)	-566,8*** (152,3)	13.85 (0.000)	0.07	96
(2) kskh	pdivhmi_aa	820,6*** (163,6)	-575,0*** (114,7)	25.13 (0.000)	0.19	58
(3) m_sec15c	pdivhmi_aa	25,7** (10,9)	-18,5** (7,7)	5.84 (0.018)	0.04	73
(4) m_tyr15	pdivhmi_aa	593,7*** (222,0)	-436,1*** (156,8)	7.74 (0.007)	0.09	73
(5) m_tyr1564	pdivhmi_aa	649,9*** (237,7)	-475,9*** (167,8)	8.04 (0.006)	0.09	73
(6) cognitive	pdivhmi_aa	434,8*** (61,4)	-307,4*** (44,1)	48.53 (0.000)	0.23	57
(7) lowsec	pdivhmi_aa	430,1*** (63,9)	-303,9*** (45,9)	43.93 (0.000)	0.20	57
(8) basic	pdivhmi_aa	160,7*** (22,6)	-113,3*** (16,3)	48.18 (0.000)	0.24	57
(9) top	pdivhmi_aa	29,8*** (7,6)	-21,2*** (5,4)	15.72 (0.000)	0.18	57
(10) school**x**cogn.	pdivhmi_aa	7722*** (2544)	-5456,0*** (1803)	9.16 (0.004)	0.11	49
(11) school**x**lowsec	pdivhmi_aa	7762*** (2559)	-5485,1*** (1813)	9.16 (0.004)	0.11	49
(12) school**x**basic	pdivhmi_aa	1776*** (478,6)	-1251,9*** (339,5)	13.60 (0.000)	0.17	49
(13) school**x**top	pdivhmi_aa	233,7*** (73,3)	-165,2*** (51,7)	10.20 (0.003)	0.14	49

Note: Dependent Variables - (1) Years of schooling; (2) Years of schooling$*$Interpersonal Trust; (3) % of population aged 15 or over with complete secondary education; (4) Years of schooling of population 15 and over, whether studying or not; (5) Years of schooling of population 15-64 who are not studying; (6) Average test score in math and science, primary through end of secondary school, all years (scaled to PISA scale divided by 100); (7) Average test score in math and science, only lower secondary, all years (scaled to PISA scale divided by 100); (8) Share of students reaching basic literacy (based on average test scores in math and science, primary through end of secondary school, all years); (9) Share of top-performing students (based on average test scores in math and science, primary through end of secondary school, all years); (10) Years of schooling$*$Average test score in math and science, primary through end of secondary school; (11) Years of schooling$*$Average test score in math and science, only lower secondary; (12) Years of schooling$*$Share of students reaching basic literacy (based on average test scores in math and science, primary through end of secondary school; (13) Years of schooling$*$Share of top-performing students (based on average test scores in math and science, primary through end of secondary school, all years).

* means significant at 10% level, ** means significant at 5% level and *** means significant at 1% level. Heteroscedasticity robust standard-errors are reported in parentheses. pdiv_aa means Predicted genetic diversity ancestry adjusted and pdivhmi_aa means Mobility index-predicted genetic diversity ancestry adjusted.

2.4.3 Results: Other Controls

We now test the robustness of the empirical relationship we described earlier to the introduction of several different controls. This exercise seeks to reduce the omitted variable bias that could be affecting our previous estimations. We introduced in a (not-shown) regression for School – as the dependent variable – all the covariates that Ashraf and Galor (2013) tested for income (see their Table 7). From those we selected the statistically significant coefficients (at the 10% level).[8] Selected covariates were social infrastructure and the percentage of population at risk of contracting malaria. These are in fact the most related variables to the expropriation risk of returns from human capital. Social infrastructure intends to measure wedges between private and social returns (this variable was originally used by Hall and Jones 1999). The higher the wedges, the less incentive to accumulate human capital. Moreover, it can be regarded as a proxy for social and family background, which has been found to influence educational achievements (see e.g. Hanushek and Woessmann, 2011).

Social Infrastructure is measured by an index of government anti-diversion policies (GADP) created from data assembled by a firm that specializes in providing assessments of risk to international investors, Political Risk Services. It is a composed indicator of law and order, bureaucratic quality, corruption, risk of expropriation, and government repudiation of contracts. The second element of Social Infrastructure captures the extent to which a country is open to international trade. This may also be related to human capital investment, as it can measure the extent to which human capital is subject to competition from human capital producing goods abroad. Moreover, the risk of contracting malaria (originally from Gallup and Sachs, 2001) is directly related to disability and death thus providing a direct effect of expropriation of potential returns from human capital.

In Table 2.4 we note that even though these two new covariates are highly statistically significant in explaining the distribution of human capital across countries, with a positive effect of Social Infrastructure and a negative effect of the percentage of population at risk of contracting malaria, the hump-shaped relationship between the different measures of human capital and genetic diversity is maintained with high statistical significance for most of human capital variables, with the exceptions of the percentage of population with secondary school (*m_sec15c*) and years of schooling in population above 15 years old (*m_tyr15* and *m_tyr1564*). With *m_sec15c* there is a marginally significant linear effect (when *pdiv_aa* squared is dropped). Results that use the alternative measure of genetic diversity (predicted through the human mobility index - *pdivhmi_aa*) are in Table A.2 in the Appendix and conclusions are in line with the previous ones. It is interesting to note that the quantitative effect of genetic diversity is now less than what emerged from the restricted regressions. In fact, a 1 percentage point increase in the lowest genetic diversity would imply 0.48 years (nearly 6 months) more in schooling (and 0.42 additional points in the school variable) and would imply nearly 3.2 months less schooling (and 0.32 points less in cognitive score) if the country were to depart from the highest levels of genetic diversity. The higher effects of increased diversity for lower levels than for higher levels is now slightly increased.

[8]Results are available upon request. With this strategy we guarantee that we tested a number of available potential macro-level determinants of human capital, thus reducing the scope for the omitted variable bias. However, the inclusion of non-significant variables would not change our main results.

Table 2.4: Human Capital and Genetic Diversity (other controls)

	Predicted genetic diversity	Predicted genetic diversity square	Social infrastructure	% of population at risk of contracting malaria	$Adj.\ R^2$ Observations
(1)	364.64**	-252.45**	5.34***	-2.6***	0.63
	(166.84)	(119.16)	(0.83)	(0.55)	100
(2)	316.29*	-217.81*	3.64***	-0.69**	0.50
	(161.02)	(114.55)	(0.61)	(0.27)	61
(3)	6.34	-4.06	0.25***	-0.1***	0.49
	(10.69)	(7.61)	(0.05)	(0.03)	91
(4)	189.2	-129.92	6.65***	-2.71***	0.70
	(189.09)	(135.25)	(0.85)	(0.58)	91
(5)	197.30	-134.29	7.24***	-2.81***	0.71
	(196.92)	(140.84)	(0.89)	(0.61)	91
(6)	361.10***	-253.80***	1.17***	-0.41	0.52
	(104.54)	(74.14)	(0.28)	(0.28)	57
(7)	371.53***	-261.49***	1.19***	-0.54	0.49
	(108.75)	(77.13)	(0.30)	(0.35)	57
(8)	147.11***	-103.07***	0.39***	-0.09	0.47
	(44.96)	(31.94)	(0.11)	(0.11)	57
(9)	24.93***	-17.71***	0.08***	-0.05**	0.44
	(7.49)	(5.36)	(0.02)	(0.02)	57
(10)	4991.90**	-3509.42**	25.90***	-19.98***	0.48
	(2288.12)	(1630.76)	(5.88)	(6.58)	53
(11)	5113.63**	-3597.50**	26.05***	-20.62***	0.49
	(2298.49)	(1637.63)	(5.92)	(6.53)	53
(12)	1256.36***	-879.96***	5.64***	-3.57***	0.54
	(440.16)	(313.88)	(1.14)	(1.19)	53
(13)	192.82***	-136.49***	0.72***	-0.55**	0.45
	(66.63)	(47.54)	(0.18)	(0.22)	53

Note: Dependent Variables - (1) Years of schooling; (2) Years of schooling$*$Interpersonal Trust; (3) % of population aged 15 or over with complete secondary education; (4) Years of schooling of population 15 and over, whether studying or not; (5) Years of schooling of population 15-64 who are not studying; (6) Average test score in math and science, primary through end of secondary school, all years (scaled to PISA scale divided by 100); (7) Average test score in math and science, only lower secondary, all years (scaled to PISA scale divided by 100); (8) Share of students reaching basic literacy (based on average test scores in math and science, primary through end of secondary school, all years); (9) Share of top-performing students (based on average test scores in math and science, primary through end of secondary school, all years); (10) Years of schooling$*$Average test score in math and science, primary through end of secondary school; (11) Years of schooling$*$Average test score in math and science, only lower secondary; (12) Years of schooling$*$Share of students reaching basic literacy (based on average test scores in math and science, primary through end of secondary school; (13) Years of schooling$*$Share of top-performing students (based on average test scores in math and science, primary through end of secondary school, all years). F-tests for the Predict Genetic Diversity Squared coefficient (not shown) always rejects for significant coefficients. Results are available upon request.

Level of significance: *** for p-value<0.01; **for p-value<0.05;* for p-value<0.1.

Values between parentheses are standard errors.

2.4.4 Robustness

One of the robustness tests provided by Ashraf and Galor (2013) is the inclusion of continent dummies. These continent dummies can account for determinants of human capital other than genetic diversity and those introduced in previous regressions, which may be continent-specific. Examples of continent-specific effects can be broad cultural heritage and (some features of) climate. Broadly speaking, we still obtain statistically significant hump-shaped effects of genetic diversity and human capital.[9] We followed the strategy described above to select a meaningful specification that may decrease the potential bias due to omitted variables. Departing from an initial regression for school, we tested the significance of all the covariates also used by Ashraf and Galor (2013), introducing continent dummies into the regression. We then selected variables with significant coefficients (at the 10% level), maintaining continent dummies. One of the significant variables is now different from above: the percentage of population at risk of contracting malaria is now replaced by the percentage of population living in tropical zones.

Almost all the human capital variables maintain the hump-shaped robust relationship with genetic diversity. In fact, when compared with the conclusions drawn from the results in Table 2.4, the differences are the following: the variables linked with the years of education for population over age 15 are now linearly (and positively) related to genetic diversity, with 5% and 10% statistical significance, meaning a small effect of nearly two months of schooling due to an increase of 1 percentage point in genetic diversity. Moreover, the hump-shaped robust (and also the linear one) relationships with three of the quality-adjusted human capital measures (for tests scores, cognitive and lowsec - and for social capital, kskh) disappears with the introduction of continent dummies. The results with the human mobility index adjustment for genetic diversity confirm all the results described above but recovers the highly significant (at 5%) hump-shaped relationship between genetic diversity and the social capital adjusted measure of human capital. Generally, for these regressions with continent dummies, a 1 percentage point increase in the lowest level of diversity implies an increase in schooling of nearly 7 months (and of 0.31 in test scores) and a 1 percentage point diversity increase in the highest level would imply a decrease of 2.3 months in schooling (0.26 in test scores).[10]

To further address robustness of our main result, we seek a causal relationship between genetic diversity and human capital. In fact, genetic diversity could also be an endogenous outcome of geographic areas with more human capital, as genetic diversity could have been improved through migration from less knowledge-endowed areas to more knowledge-endowed areas or the other way around (specific examples of such migrations were the barbarian invasions of the Roman Empire and the colonization of Africa). On the contrary, well endowed areas were also able to erect barriers to deter colonization (e.g. China's Great Wall and Roman Legions for centuries). The issue is that these types of migrations did in fact increase genetic diversity, then genetic diversity and human capital may be outcomes of some other determinants of development, and thus genetic diversity could not be regarded as exogenous towards human capital today. To minimize the endogeneity problem, Ashraf and Galor (2013) constructed a measure of predicted genetic diversity based on "physical" distance from East Africa, which was the variable we have been using in our paper. As the authors put it "given the obvious exogeneity of migratory distance from East Africa with respect to development outcomes in the Common Era, the use of migratory distance to project genetic diversity alleviates concerns regarding

[9] Results are available upon request.

[10] These quantitative effects are slightly decreased when using human mobility index adjustment for genetic diversity.

the potential endogeneity between observed genetic diversity and economic development (...) Specifically, the identifying assumption being employed here is that distances along prehistoric human migration routes from Africa have no direct effect on economic development during the Common Era." (Ashraf and Galor, 2013: 6, 14).

However, we want to document further that the hump-shaped relationship between human capital and genetic diversity can be regarded as a causal relationship. Here we avoided the introduction of more controls, because their own exogeneity might be at stake (e.g. the social infrastructure measure).

There are two main issues with validity of IV estimates. One is the possibility of weak instruments, that is, instruments that are not sufficiently correlated with the instrumented variables. The other is the adequacy of instruments, i.e., their potential correlation with the error term. We accounted for both problems in our estimations. First, we have carefully analyzed all the first-step regressions for high significance of regressors. Second, we have analyzed the result of the Kleibergen-Paap rk LM statistic for under-identification (insufficient instruments), and the Cragg-Donald Wald F statistic for weak-identification (weak instruments). We also used the Stock-Wright LM S statistic for the joint significance of endogenous variables in the main regression, which is a test of robust inference, robust to weak instruments. If this test is rejected it means that instruments can be used to explain the dependent variable, not only through the instrumented variable. In our case this would mean that distance could be used to explain human capital (see the discussion above about the exclusion of migratory distance as a predictor of human capital development several centuries later). Third, we performed the Hansan J-statistic to test for adequate instruments. For a good IV regression, all the tests but the Stock-Wright LM S statistic and J-statistic should be rejected. Fourth, we have applied an endogeneity test, which indicates whether we can treat genetic diversity as exogenous in the context of the regressed equations. We strictly followed the instrumentation technique used by Ashraf and Galor (2013), and obtain a statistically significant hump-shaped influence of genetic diversity in human capital in which all the other predictors of human capital are also statistically significant. Overall, the tests indicate that we can reject the hypothesis that instruments are weak and do not reject that instruments are adequate.

Table 2.5 presents examples of regressions for the enlarged sample in which we used *predicted* genetic diversity to predict human capital. The results reveal the same robust relationship, and the tests indicate that instruments are not weak and are not correlated to the error term, which clearly indicates that inference can be made. In fact, the endogeneity test indicates that the predicted genetic diversity could indeed have been treated as exogenous. It is also worth noting that quantitatively the coefficients that define the hump-shaped relationship between human capital and genetic diversity are not very different from the coefficients estimated by OLS earlier in the paper. It is striking that the threshold level above which genetic diversity causes human capital to diminish is almost the same as before, 0.7!

We have experimented with several specifications for the different human capital variables we studied above and from all experiments it was possible to demonstrate that genetic diversity can be regarded as *causing* human capital to rise for low levels of diversity and as *causing* human capital to decrease for high values of genetic diversity.[11]

[11] We ran several regressions. Almost all the results suggest that predicted genetic diversity (ancestrally adjusted) can be treated as exogenous within regressions that use the same instrument set as in the columns of Table 2.5 - exceptions are only for years of schooling, for population above 15 years old, and for the share of students with top scores in tests (m_tyr15, m_tyr1564 and top). Despite that, even for the exceptions, it is possible to find instrument sets that allow for the endogeneity test to fail rejection

Table 2.5: Human Capital and Genetic Diversity (2SLS estimates)

	(1)	(2)	(3)
Dependent Var.	School	Cognitive	Cognitive $\times$School
Predicted Diversity	873.0** (0.014)	557.4*** (0.000)	8008** (0.041)
Predicted Diversity Square	-614.6** (0.013)	-394.5*** (0.000)	-5674** (0.041)
Kleibergen-Paap rk LM statistic	31.6*** (0.000)	16.0** (0.026)	14.7** (0.041)
Cragg-Donald Wald F statistic	$8.9^{\dagger}$ (6.20)	$9.8^{\dagger}$ (6.20)	$9.7^{\dagger}$ (6.20)
Stock-Wright LM S statistic	12.9 (0.116)	15.3* (0.054)	7.8 (0.453)
Hansen J-Statistic	6.91 (0.330)	4.26 (0.119)	7.25 (0.299)
Endog. Test	2.49 (0.288)	3.79 (0.150)	3.40 (0.183)
N	123	71	63

Note: Excluded Instruments: aerial distance from East Africa, aerial distance from East Africa (square), terrestrial distance from London, Tokyo, and Mexico, terrestrial distance from London, Tokyo, and Mexico (square). Kleibergen-Paap rk LM statistic tests underidentification, under the null of the matrix of reduced form coefficients has rank=K1-1. Cragg-Donald Wald F statistic tests the null under which equation is weakly identified. This is compared with the Stock-Yogo weak ID test critical values, which are reported in parentheses in that line. Stock-Wright LM S statistic tests the null under which the joint endogenous regressors have null coefficients. Hansen J-Statistic tests the null under which the instruments are valid, i.e., uncorrelated with the error term. Level of significance: *** for p-value<0.01; **for p-value<0.05;* for p-value<0.1. † 6.20 is the critical value for relative IV bias of 20% (of the OLS bias). Values inside parentheses are p-values, except for the Cragg-Donald Wald F statistic. All regressions include continent dummies.

2.5 Conclusion

We investigate the relationship between human capital and the ancestral genetic diversity of populations. The paper highlights a new channel through which genetic diversity can affect development, through human capital. We have devised a very simple model in which human capital benefits from an increasing (inherited) variety of genetic traits (heterozygosity), which enhance learning abilities. Additionally, a cost of human capital which depends increasingly on genetic diversity is essential to depict a hump-shaped relationship between genetic diversity and the human capital supply. This cost represents the additional effort economic agents have to support in order to overcome the negative influence of very diverse genetic backgrounds on the school environment. Despite its simplicity, the model encompasses the interplay between nature and nurture in the human capital supply of the economy, the presence of inherited family genetic traits and the costs of diversity on learning environment.

We based our empirical study on a database of human capital variables coming from Cohen and Soto (2007) - for measures of quantity of human capital and from Hanushek and Woessmann, 2012 - for measures of quality of human capital and then merged it with the database of genetic diversity, from Ashraf and Galor (2013). We found a hump-shaped relationship between human capital and genetic diversity, confirming the idea that the influence of genetic diversity on development may be through human capital. A 1% change in low levels of genetic diversity may imply large effects in schooling that can oscillate between more 4 and 12[12] months of schooling (more 0.31 to 0.48 points in international tests scores) and negative effects when there is high genetic diversity (less 2 to 10 months of schooling and near less 0.3 to 0.4 points on scores).

We show and discuss a number of robustness tests with instrumental variables regressions. The overall conclusion is that the hump-shaped relationship between human capital and genetic diversity can indeed be regarded as a causal relationship. Thus, human capital outcomes may

(accepting that instrumented variables can be treated as exogenous). Moreover, most variables linked with human capital are robustly related to the predicted genetic diversity (ancestrally adjusted) - in the typical hump-shaped manner in regressions specified as those in Tables 4 and 5. More important than that, for those variables that could not be statistically treated as exogenous (the three mentioned above), it is possible to specify well-behaved IV models that highlight the robust hump-shaped relationship with predicted genetic diversity.

[12] Using the most preferable specifications (Tables 2.4 and 2.5 and regressions with controls and continent dummies).

have been set onto their current paths millenia ago, when great human migrations shaped the countries' genetic diversity that we see today.

Chapter 3

Technology in 1500 and Genetic Diversity

3.1 Introduction

The literature on the deeply-rooted determinants of development has been greatly developed in the last decade. Hall and Jones (1999) demonstrated the importance of social infrastructure, a composed measure of law of rule and other institutional measures, which was then followed by Glaeser *et al.* (2004). It is interesting to note that Glaeser's paper highlights the role of human capital as a more significant source of growth than institutions but does not show any causal relationship between institutions and human capital. Rodrick *et al.* (2004) showed evidence for the supremacy of institutions related to geography and integration as determinants of economic development. However, Olsson and Hibbs Jr. (2005) showed the strong influence of geographical and biogeographical factors in determining the current level of development. In related literature on institutions, Sokoloff and Engerman (2000) and Acemoglu *et al.* (2005) have stressed the role of colonialism, while the effects of ethnolinguistic fractionalization were examined by Easterly and Levine (1997) and Alesina *et al.* (2003).[1] Moreover, the historical impact of sociocultural factors has been highlighted by Barro and Mccleary (2003), Tabellini (2008), and Guiso *et al.* (2009).

Nevertheless, the study of the deeply-rooted determinants of technology and innovation is a very recent strand of the literature. Despite of the modern common wisdom among economic historians that the early industrial revolution (in England) was driven by macroinventions which are considered mostly exogenous and were not determined by previous investment in education (Mitch, 1993; Mokyr, 1990; Galor, 2005; Galor and Moav, 2006), the research about the initial conditions that drove the first industrial revolution is not abundant (see e.g. Gómez and Sequeira, 2012). Deeply rooted determinants of technological advance can be understood as some of those very initial conditions. Genetic distance proved to be relevant for technology adoption as Spolaore and Wacziarg (2012) found on their empirical study; populations that have greater genetic distance from innovator have lower aggregate TFP (Total Factor Productivity), more technology usage lags, less absorption on extensive margin (technology adoption which captures the portion of potential adopters that actually adopted the technology) of technology adoption and lower rates of technology usage at the level of disaggregated technologies. Another important determinant of technology diffusion are lobbies; Comin and Hobijn (2009) with their empirical study for over 20 technologies and 23 of the world's leading industrial economies identify specific effects of institutions that proved to strongly affect technology diffusion. In particular, when institutions try to generate more flexibility of the legislative authority to accelerate the speed of diffusion it actually is making it easier for lobbies to induce him to promote political barriers that slows down the technologies diffusion.[2] For Comin et al. (2013) spatial

[1]The influence of genetic diversity on ethnolinguistic fractionalization has been studied by Ahlerup and Olsson (2012).

[2]Most of the work of Diego Comin and co-authors is supported by the database CHAT (Comin and Hobjin, 2009), which is a cross-country historical database of adoption of technology that contains more than 100 technologies for more than 150 countries since 1800. The database is available at

distance is determinant for diffusion, countries that are far away from adoption leaders have a slowly adoption but the spatial effects vanish over time. Focusing on agricultural societies Baker (2008), using OLS and 2SLS methods and with a different approach of independent variable technological sophistication (with 2 different measures, first measures the degree that a society is specialized in performing metal working, pottery making and leather working, second measures the sophistication of writing and record-keeping system in each society) - affirms that growth effects are important to determine the technological sophistication giving some support to assumption that the adoption of agriculture changed the nature of endogenous growth dynamics. Apparently technologies adoption has an historical effect on countries following Comin and Mestieri (2013). Adoption lags (time lag that a country takes between technology invention and the adoption) have converged across countries over the last 200 years but penetration rates (difference between countries adoption) have diverged, the evolution of adoption patterns was accompanied by evolution of income growth across countries and, therefore, explaining the divergence and convergence between countries for over the last two centuries. Complementing these results, Comin et al. (2006) looked for patterns on technology diffusion; they found that taking into account the typical technology there has been convergence at an average rate of 4 percent per year and the speed of convergence for technologies since 1925 has been three times higher than before 1925 even when separating countries by being members, or not, of OECD. Ashraf and Galor's (2013) work is probably one of the most influential recent papers in the field of economic development.[3] Their paper reports a significant relationship between genetic diversity determined ancestrally and economic development in the present. The paper illustrates the relationship between genetic diversity and development through the positive and decreasing effect of genetic diversity on technology and a negative effect of genetic diversity on output representing inefficiency in production. The hump-shaped empirical relationship exposed by the authors builds on two opposite effects of genetic diversity. First, an increase in diversity enhances production possibilities as a wider spectrum of traits is more likely to contain those that are complementary to the advancement of superior technologies. In fact, some competition for survival, as natural selection explains, also increases adaptability and improves the society's ability to successfully introduce new and better technologies. However, after a certain level of genetic diversity, further rise would increase the scope for disarray and mistrust, increasing the probability of conflict. Ashraf and Galor (2013) focus their paper on explaining development in 1500 due to genetic diversity. In their paper the technological development of countries in 1500 is measured by the population density. In our paper, we change the dependent variable and use instead the technologies measures of Comin *et al.* (2010) in 1500. We aim to provide a more direct measure of technology than population density and also provide some robustness test of the initial result as we use different technologies measures. Comin *et al.* (2010) created a database on technologies adoption in different fields. The database covers the technology usage (or not) in each country (in 1500) for agriculture, transportation, communication, industry, military and an average sector (the average of previous sectoral indexes). The remaining of the paper is organized as follows. In section 3.2 we describe the data. In section 3.3 we present the main results, subject to a number of extensions. In section 3.4 we discuss results on robustness tests linked with the introduction of continent dummies and instrumental variables estimations. Section 3.5 concludes.

http://www.nber.org/data/chat.

[3]The paper was a lead article in American Economic Review and was the science's editor choice in September 2012.

Table 3.1: Technology Variables

Classifications of technologies in 1500 (tech)	Name	Technologies included in each classification
total/Average	avr	Average technology adoption (an average of agriculture, transportation, communication, industry and military).
Agriculture	agr	Hunting & Gathering; Pastoralism; Hand Cultivation; Plough Cultivation.
Transportation	tra	Ships capable of crossing the Atlantic/Pacific/Indian oceans; Wheel; Magnetic compass; Horse powered vehicles.
communication	com	Movable block printing; Woodblock or block printing; Books; Paper.
Industry	ind	Steel; Iron.
Military	mil	Standing army; cavalry; Firearms; Muskets; Field artillery; Warfare capable ships; Heavy naval guns; Ships (+180 guns), +1500 ton Deadweight.

Note: for more details refer to table 3 in Comin et. al. (2010).

Table 3.2: Descriptive Statistics

Variable	Mean	Std. Dev.	Min	Max
(a) pdiv	.7267	.0269	.6279	.7743
(b) pdivhmi	.7229	.02904	.6178	.7826
(c) adiv	4.8623	2.8126	.4089	10.8622
(1) avr	.4640086	.3142213	0	1
(2) agr	.715942	.3066625	0	1
(3) tra	.2434783	.2679128	0	1
(4) com	.4202586	.3936992	0	1
(5) ind	.573913	.4148741	0	1
(6) mil	.356681	.379593	0	1

Notes: (a) Predicted genetic diversity; (b) Mobility index-predicted genetic diversity; (c) Observed genetic diversity. Details are given in the appendix f of Ashraf and Galor (2013).; (1)-(6) are defined in the previous table.

3.2 Data and Sources

Explanatory variables rely on the database from Ashraf and Galor (2013) who include variables that measure genetic diversity, as adjusted to migratory movements as well as other explanatory variables (neolithic transition time, % of arable land, latitude, suitability of land for agriculture). For our benchmark analysis we use the predicted genetic diversity, and the mobility adjusted predicted genetic diversity for 1500 as the main explanatory variables.[4]
Table 3.1 summarizes the dependent variables used in this paper (from Comin *et. al.*, 2010) – Classifications of technologies in 1500. These variables are average measures of usage of certain technologies (e.g. The Wheel, Magnetic Compass, Steel) in 1500. Most technologies adopt the values 1 or 0 according to which they were used or not in each country. Then, each of the classifications (agriculture, transportation, communication, industry and military and an average over the previous) are averages of the different technologies usage for each classification. There are those classifications that we use as dependent variables in our regressions. Table 3.2 reports descriptive statistics for the dependent and explanatory variables.

3.3 Results

We relate each of the variables for technology adoption in 1500 (as dependent variables) with genetic diversity, using as controls all the variables used by Ashraf and Galor (2013) in their benchmark regressions for 1500. Tables 3.3 and 3.4 show the results. All regressions reveal an impressive robust relationship between genetic diversity and our proxies for technological adop-

[4]The largest list of countries used (106) is detailed in the appendix. Detailed lists of countries for each regression are available upon request.

tion. They also draw a hump-shaped relationship between genetic diversity and technological adoption. The cut-off point above which the marginal effect of increases in genetic diversity begins to be negative is around 0.7. An increase of 1 percentage point (0.01) in genetic diversity for countries with the lowest genetic diversity would imply an increase in 0.043 score for agriculture, 0.050 for communications, 0.028 for transportation, 0.078 for industry, 0.035 for military and 0.047 for the average sector. These are sizeable effects given that the technological adoption proxies are measured between 0 and 1, corresponding to values between 2.8% (transportation) to 7.8% (industry).
For those countries with the highest genetic diversity, a further increase of 0.01 in diversity would cause decreases in technology from 0.02 (2.1%) in transportation to 0.09 (8.6%) in agriculture. There is evidence of a relatively greater positive effect of diversity on technology for all sectors, except for agriculture, in which the negative effect of increasing genetic diversity is greater for high levels of diversity than the positive effect that occurs in low levels of genetic diversity. The only exceptions for the hump-shaped robust relationship are for transportation and military associated technologies when genetic diversity is adjusted with the human capital mobility index (Table 3.4). However, in both of those cases it can be shown that there is a robust linear relationship between genetic diversity and technological adoption, meaning that an increase of 1 percentage point in genetic diversity (for all levels of diversity) would imply increases in technological adoption linked with transportation of 1.4% and with military of 1.8%. The most significant variables in determining the distribution of these proxies of technology around the world, apart from genetic diversity, are the neolithic transition time and absolute latitude.

Table 3.3: Regressions for Technology Adoption in A.D. 1500

	Predicted Genetic Diversity	Predicted Genetic Diversity Square	log [Neolithic Transition Timing]	log [Percentage of Arable Land]	log [Absolute Latitude]	log [Land Suitability for Agriculture]	$adj.\ r^2$ / Observations
(1) agr	59.80***	-44.16***	0.29***	0.06**	-0.002	-0.004	0.55
	(16.64)	(12.20)	(0.04)	(0.03)	(0.02)	(0.03)	106
(2) com	47.10***	-33.51***	0.27***	0.05	0.15***	0.03	0.71
	(10.67)	(7.98)	(0.04)	(0.03)	(0.02)	(0.03)	106
(3) tra	23.94***	-16.84***	0.23***	0.04**	0.07***	-0.024	0.67
	(7.25)	(5.44)	(0.03)	(0.02)	(0.02)	(0.02)	106
(4) mil	30.74**	-21.63**	0.29***	0.05**	0.11***	-0.02	0.60
	(13.80)	(10.33)	(0.04)	(0.02)	(0.03)	(0.03)	106
(5) ind	57.28***	-39.43***	0.27***	0.03	0.06***	-0.02	0.74
	(14.10)	(10.45)	(0.03)	(0.03)	(0.02)	(0.03)	106
(6) avr	43.77***	-31.11***	0.27***	0.04**	0.08***	-0.01	0.78
	(8.18)	(6.09)	(0.02)	(0.02)	(0.02)	(0.02)	106

Note: Dependent variables - (1) Average technology adoption in agriculture in period 1500 CE; (2) Average technology adoption in communications in period 1500 CE; (3) Average technology adoption in transportation in period A.D. 1500; (4) Average technology adoption in military in period A.D. 1500; (5) Average technology adoption in industry in period A.D. 1500; (6) Average of the sectoral technology adoption indexes in period A.D. 1500;
Level of Significance: *** for p-value<0.01; **for p-value<0.05;* for p-value<0.1. Values between parentheses are standard errors. f-tests for the predict genetic diversity squared coefficient (not shown) always rejects for significant coefficients. Results are available upon request.

Table 3.4: Regressions for Technology Adoption in A.D. 1500

	Mobility index-predicted genetic diversity	Mobility index-predicted genetic diversity square	log [Neolithic transition timing]	log [Percentage of arable land]	log [Absolute latitude]	log [Land suitability for agriculture]	$adj.\ r^2$ / Observations
(1) agr	35.59**	-26.27**	0.37***	0.04	-0.01	0.002	0.52
	(15.41)	(11.39)	(0.05)	(0.04)	(0.02)	(0.03)	94
(2) com	32.02***	-22.57***	0.31***	0.04	0.13***	0.03	0.69
	(9.72)	(7.33)	(0.06)	(0.04)	(0.02)	(0.03)	94
(3) tra	7.73	-4.71	0.30***	0.02	0.06***	-0.01	0.67
	(6.83)	(5.17)	(0.03)	(0.02)	(0.02)	(0.02)	94
(4) mil	12.94	-8.36	0.37***	0.02	0.10***	0.003	0.60
	(10.37)	(7.84)	(0.05)	(0.03)	(0.03)	(0.03)	94
(5) ind	27.12**	-17.08*	0.34***	-0.01	0.07***	-0.002	0.71
	(11.68)	(8.81)	(0.04)	(0.03)	(0.02)	(0.02)	94
(6) avr	23.08***	-15.80***	0.34***	0.02	0.07***	0.005	0.77
	(7.74)	(5.79)	(0.03)	(0.02)	(0.02)	(0.02)	94

Note: Dependent Variables - (1) Average technology adoption in agriculture in period 1500 CE; (2) Average technology adoption in communications in period 1500 CE; (3) Average technology adoption in transportation in period A.D. 1500; (4) Average technology adoption in military in period A.D. 1500; (5) Average technology adoption in industry in period A.D. 1500; (6) Average of the sectoral technology adoption indexes in period A.D. 1500;
Level of Significance: *** for p-value<0.01; **for p-value<0.05;* for p-value<0.1. Values between parentheses are standard errors. f-tests for the predict genetic diversity squared coefficient (not shown) always rejects for significant coefficients. results are available upon request.

3.4 Robustness

One of the tests provided by Ashraf and Galor (2013) is the inclusion of continent dummies. These continent dummies can account for determinants of technological adoption other than genetic diversity and the those introduced in previous regressions, which may be continent-specific. Examples of continent-specific effects can be broad cultural heritage and (some features of) climate. Broadly speaking, we still obtain statistically significant hump-shaped effects of genetic diversity and technology adoption in 1500. These quantitative effects are slightly decreased when using human mobility index adjustment for genetic diversity, a feature that is also clear by the previous comparison of results on Table 3.3 with results on Table 3.4.[5]
To further address robustness of our main result, we seek a causal relationship between genetic diversity and technological adoption. In fact, genetic diversity could also be an endogenous outcome of geographic areas with more technological adoption, as genetic diversity could have been improved through migration from low technology areas to high-technology areas (e.g. the barbarian invasions of the roman empire). On the contrary, high technology areas were also able to erect barriers to deter colonization (e.g. China's great wall and roman legions for centuries). The issue is that these types of migrations did in fact increase genetic diversity, then genetic diversity and technological adoption may be outcomes of some other determinants of development, and thus genetic diversity could not be regarded as exogenous towards technological adoption in 1500. To minimize the endogeneity problem, Ashraf and Galor (2013) constructed a measure of predicted genetic diversity based on "physical" distance from East Africa, which was the variable we have been using in our paper. As the authors put it "given the obvious exogeneity of migratory distance from East Africa with respect to development outcomes in the common era, the use of migratory distance to project genetic diversity alleviates concerns regarding the potential endogeneity between observed genetic diversity and economic development (...) specifically, the identifying assumption being employed here is that distances along prehistoric human migration routes from Africa have no direct effect on economic development during the common era." (Ashraf and Galor, 2013: 6, 14). Following what was done by Ashraf and Galor (2013) to test exogeneity toward development (the log of population density), we tested the exogeneity of *observed* genetic diversity in explaining technological adoption around the world. First, we have regressed the average sectoral technological index in 1500 (*avr*) on observed genetic diversity (linear and square terms) migratory distances from East Africa (linear and square terms) together with log neolithic transition time, log percentage of arable land, log absolute latitude, log land suitability for agriculture and continental dummies. The linear and squared terms of observed genetic diversity are significant at 5% and both terms on distance are non-significant. We repeat the procedure substituting distances with the mobility index (linear and squared terms). This means that when considering genetic diversity and distances or migrations as potentially simultaneous determinants of technological development, we end out concluding that genetic diversity is always statistically stronger as a candidate to explain technological adoption.
Secondly, we replicated 2SLS regression those authors have presented in their Table 2, column (6) but now with the average sectoral technological index in 1500 (*avr*) we have used earlier, as the dependent variable. This serves as a test for exogeneity of *observed* genetic diversity for the smaller sample. Given that the *observed* genetic diversity could be established as exogenous, then the predicted genetic diversity would be reasonably considered as exogenous by the arguments exposed earlier. This will be column (1) in our table 3.5.

[5]Results are available upon request.

However, we want to document further that the hump-shaped relationship between technological adoption and genetic diversity in 1500 can be regarded as a causal relationship. Thus, we used the geographical aerial distances to East Africa, terrestrial distances to London and Mexico as instruments to the predicted genetic diversity for the larger sample. In fact, geographical distances to given points in the word, despite being highly correlated with predicted genetic diversity (remember that terrestrial distances to East Africa were used to construct predicted genetic diversity, departing from observed genetic diversity), they are not reasonably assumed to be, by themselves, direct determinants of technological adoption in 1500, many centuries after prehistoric migration routes. Thus, geographical distances are good candidates to instrument the predicted genetic diversity. Columns (2) and (3) present 2SLS regressions for the average sectoral technological index in 1500 (*avr*). Note that we introduce a set of covariates that are arguably exogenous.

There are two main issues with validity of IV estimates. One is the possibility of weak instruments, that is, instruments that are not sufficiently correlated with the instrumented variables. The other is the adequacy of instruments, e.g., their potential correlation with the error term. We accounted for both problems in our estimations. First, we have carefully analyzed all the first-step regressions for high significance of regressors. Second, we have analyzed the result of the Kleibergen-Paap rk lm statistic for under-identification (insufficient instruments), and the Cragg-Donald Wald F statistic for weak-identification (weak instruments). We also used the Stock-Wright lm s statistic for the joint significance of endogenous variables in the main regression, which is a test of robust inference, robust to weak instruments. If this test is rejected it means that instruments can be used to explain the dependent variable, not only through the instrumented variable. In our case this would mean that distance could be used to explain technological adoption (see the discussion above about the exclusion of migratory distance as a predictor of technological adoption several centuries later). Third, we performed the Hansan J-statistic to test for adequate instruments. For a good IV regression, all the tests but the Stock-Wright lm S statistic and J-statistic should be rejected. Fourth, we have applied an endogeneity test, which indicates whether we can treat genetic diversity as exogenous in the context of the regressed equations. Column (1) shows us a regression for the restricted sample. It used the *observed* genetic diversity (used by Ashraf and Galor (2013) to predict genetic diversity in the enlarged country sample). We strictly followed the instrumentation technique used by those authors, and obtain a statistically significant hump-shaped influence of genetic diversity in technology in which all the other predictors of technological development are also statistically significant. The tests indicate that we can reject the hypothesis that instruments are weak. Columns (2) and (3) present regressions for the enlarged sample in which we used *predicted* genetic diversity to predict technological development in 1500. The results reveal the same robust relationship, indicating that instruments are not weak and are not correlated to the error term. Thus, inference can be made. However, contrary to what happens in column (1), the endogeneity test now indicates that statistically *predicted* genetic diversity cannot be treated as exogenous. It is also worth noting that quantitatively the coefficients that define the hump-shaped relationship between technological development and genetic diversity are not very different from the coefficients estimated by OLS earlier in the paper, with the greater quantitative difference being those of coefficients in columns (2) and (3). It is striking that the threshold level above which genetic diversity causes technological adoption to diminish is almost the same as before, 0.7!

Table 3.5: Technological Adoption and Genetic Diversity (2SLS Estimates)

Dependent Var.	(1) *avr*	(2) *avr*	(3) *avr*	(4) *avr*	(5) *avr*
Observed Diversity	42.1** (0.048)	-	-	-	-
Observed Diversity Square	-34.8* (0.051)	-	-	-	-
Predicted Diversity	-	29.0*** (0.001)	28.8*** (0.001)	-	-
Predicted Diversity Square	-	-21.0*** (0.001)	-20.9*** (0.001)	-	-
Mobility Index-predicted Genetic Diversity	-	-	-	20.2*** (0.002)	19.3*** (0.003)
Mobility Index-predicted Genetic Diversity Square	-	-	-	-14.72*** (0.004)	-14.01*** (0.006)
log Neolithic Transition Time	0.22*** (0.004)	0.22*** (0.000)	0.22*** (0.000)	0.25*** (0.000)	0.25*** (0.000)
log Percentage Arable Land	0.14*** (0.003)	0.05*** (0.006)	0.05*** (0.006)	0.03 (0.133)	0.03 (0.133)
log Absolute Latitude	0.06*** (0.003)	0.03* (0.056)	0.01* (0.056)	0.02 (0.108)	0.02 (0.109)
log Land Suit. for Agriculture	0.06* (0.051)	0.03* (0.079)	0.02* (0.076)	-0.02 (0.197)	-0.02 (0.204)
Kleibergen-Paap rk lm Statistic	-	41.2*** (0.000)	41.7*** (0.000)	30.3*** (0.000)	29.5*** (0.000)
Cragg-Donald Wald F Statistic	4.6^{δ}	$446.8^{\dagger\dagger\dagger}$ (11.04)	$353.7^{\dagger\dagger\dagger}$ (13.97)	$152^{\dagger\dagger\dagger}$ (11.04)	$145.48^{\dagger\dagger\dagger}$ (13.97)
Stock-Wright lm s Statistic	2.8 (0.242)	8.8* (0.066)	8.9 (0.112)	14.02*** (0.007)	14.41** (0.013)
Hansen J-Statistic	exact. id.	1.57 (0.457)	1.59 (0.662)	3.23 (0.199)	4.07 (0.254)
Endog. Test	3.16 (0.206)	10.68*** (0.005)	16.88*** (0.000)	6.260** (0.044)	6.05** (0.049)
N	19	106	106	94	94

Note: Excluded instruments - (1) 2 instruments: first-stage fitted values of observed genetic diversity (square) and migratory distance from east Africa, equation is exactly identified; (2) and (4) 4 instruments: aerial distance from east Africa, aerial distance from east Africa (square), terrestrial distance from London and terrestrial distance from London (square); (3) and (5) 5 instruments: the same as in column (2) and (4) plus geodesic centroid latitude; Kleibergen-Paap rk lm statistic tests underidentification, under the null of the matrix of reduced form coefficients has rank=k1-1. Cragg-Donald Wald F statistic tests the null under which equation is weakly identified. this is compared with the Stock-Yogo weak id test critical values, which are reported in parentheses in that line. stock-wright lm s statistic tests the null under which the joint endogenous regressors have null coefficients. Hansen J-statistic tests the null under which the instruments are valid, i.e., uncorrelated with the error term.

Level of significance: *** for p-value<0.01; **for p-value<0.05;* for p-value<0.1. δ in column (1) 4.58 is the critical value for 15% critical iv size; † † † in columns (2) and (4), 11.04 is the critical value for relative IV bias of 5% (of the OLS bias); in columns (3) and (5), 13.97 is the critical value for relative IV bias of 5% (of the OLS bias).

values inside parentheses are p-values, except for the Cragg-Donald Wald F statistic. All regressions include continent dummies.

We ran several alternative regressions. Even though in the regression for the average sectoral technological development the predicted diversity does not appear to be statistically exogenous, for the technology associated with agriculture, communications, and industry, the endogeneity test is not rejected. For these technological adoption proxies IV regression is also well-behaved regarding the instruments' properties and the hump-shaped relationship is revealed. We present those results in tables B.2.1 to B.2.5.

3.5 Conclusion

Despite the development of the literature on the deeply-rooted determinants of development, the study of the historically rooted determinants of technological development has been scarce. Ashraf and Galor's (2013) have studied the effect of genetic diversity on development using technology as the theoretical link between genetic diversity and development. That article argued that an increase in diversity enhances production possibilities as a wider spectrum of traits is more likely to contain those that are complementary to the advancement of superior technologies. However, those authors used the population density in 1500 as a proxy for technological development. Alternatively, we use direct measures of technological adoption provided by Comin *et. al.* (2010) to address such relationship. This study highlights a strong hump-shaped relationship between genetic diversity and technological developments in 1500. This means that some of the technological achievements may stem from the genetic diversity mostly determined more than a millennium ago. Results are robust to the introduction of several controls, and to IV estimation.

Chapter 4

Why Did Inventions Occur in Some Countries and not in Others?

4.1 Introduction

There is a diverse literature on the drivers of industrial revolutions which could be translated into the drivers of earlier innovation or Total Factor Productivity (TFP) growth. There are essentially three different views about the triggers of the first innovations in Britain, which shaped the beginning of the industrial revolution. First, a view that innovations occurred in Britain because it paid to invent them there. This is the view of Robert Allen (2009a, 2009b) who bases the argument on the relative factor prices in place in Britain in the middle of the Eighteen century. In fact, Britain experienced that times relatively high wages and relatively cheap capital and energy which made the incentives to innovate and substitute labor by capital. According to this view, this is the reason why the industrial revolution began in Britain and not, e.g., in France, where the labor was much cheaper. Allen also recognizes that the supply of inventors is important and he points out that in years prior to 1800, a cultural revolution had happened in Britain and that human capital is those times much greater than one or two centuries before. This relative supply of factors reasoning has been showed to create different patterns of industrial revolutions using models of endogenous growth (see e.g. Iacopetta 2010 and Gómez and Sequeira 2012).

Second, a view that justifies the British Industrial Revolution due to the better institutions that have been earlier built in Britain and not in any other part of the world. This is a view argued by Joel Mokyr (2009). According to the Mokyr's view "Britain became the leader of the Industrial Revolution because, more than any other European economy, it was able to take advantage of its endowment of human and physical resources thanks to the great synergy of the Enlightenment: the combination of the Baconian program in useful knowledge and the recognition that better institutions created better incentives" (p.122). Mokyr believes that the combination of skilled scientists, engineers, entrepreneurs and craftsman allow for the invention and adoption of technologies by firms. Thus, the Enlightenment age in Britain is viewed with two main effects: it improved technological capabilities and institutional quality. The Baconian program comprised research based on experimentation and scientific method, directing the research to solve practical problems, and making results accessible (p.40). Mokyr acknowledges that the impact of the Enlightenment on institutions is hard to quantify but argues that the success of its ideology reduced rent-seeking and promoted competitive markets (p.63). It was manifested in terms of legislation such as the abolition of the Corn Laws but also strengthened informal institutions in the form of social norms that favored gentlemanly capitalism rather than opportunistic behavior (ch.16). Clark (1996) also focused on the important political and institutional evolutions that preceded and influenced the Industrial Revolution. As the author writes: "The years between the Glorious Revolution and the Industrial Revolution saw widespread change in the British economy: the transport system was radically improved; a large scale conversion to purely private agriculture was accelerated; new institutions of finance and commerce were put in place; and the government debt was regarded as the safest asset in the economy" (p. 564). Third, a view based on the unified growth theory (due to Galor, 2005) which argues that the transition to a post-Malthusian epoch (i.e. an Industrial Revolution) was due to scale effects provided by increased or larger population. Then, the transition to sustainable growth is supported by an increasing demand for human capital, provided by families that increasingly choose quality of the offspring in opposition to quantity.

Some empirical work has been published on the growth rates and factors in the Industrial Revolution. Stockey (2001) shows the rise of industry in the composition of British GDP between 1760 and 1850. Greasley and Oxley (2007) shows that patenting rose sharply during the Industrial Revolution although they also show that this process did not cause (but is caused by) the industrialization. However, the role of technologies and of industry in the Industrial Revolution growth period has been almost consensual. Despite Britain had presented only a modest average growth rate, Broadberry et. al. (2011) clearly shows that the industrial sector was the one that presented the fastest growth (when compared with services and agriculture) between 1760 and 1860 and with accelerating growth during the period. The importance of technology has been emphasized e.g. by the work of Nicholas Crafts (see e.g. Crafts, 1992, 2004), where estimates of TFP growth has been presented and compared with those from other references.

Our work is an empirical attempt to answer the question "why technologies appeared where they appeared?". This may be interpreted as a re-statement of the question "why in England?" that many authors addressed before. However, we depart from the existing literature in two main ways. First, we study a number of documented technologies that were invented throughout the centuries (e.g. tractor, cellphone or spindle mule); second, we placed the invention in some country, thus we are not studying England only. We want to contribute to answer the question "why in England?" but also "why not in England?"

The paper is organized as follows. Section 4.2 presents the data and respective sources. Section 4.3 presents the main results and Section 4.4 concludes.

4.2 Data and Sources

We began collecting data on all the 104 technologies for which Comin et al. (2009). However, we do not use data from the CHAT dataset, we only collected the name of the technologies. Then, we searched for different sources (see Appendix A) and collected the date of the invention of all the technologies invented after 1600. There were 14 technologies invented before 1600 and 6 other for which there was not possible to identify a year of invention. So, we remain with 84 technologies to be studied. This allowed to build two different and alternative dependent variables: a dummy variable which takes the value 1 for the country which invented the given technology and 0 for all the other countries and then a variable which measure how many years ago the technology was invented (this was made subtracting the year of invention from the year 2000). For the first variable, we have 84 technologies per country (which can be an inventor or not) and 10 countries which have invented some technology after 1600. While the first variable allows to identify the reasons why a given technology was invented in a given country and not on the others, the second allows to explain the reasons why some countries were pioneer in certain technologies. Leaders in innovations were the United States of America (USA), with 40 technologies invented, followed by the United Kingdom (UK), with 17, Germany, with 15, France (4), Japan (3), Switzerland, Netherlands, Austria, Australia and Russia, with 1 each.

For the explanatory variables we tried to match the alternative explanations that previous literature have identified as drivers of the Industrial Revolution. The shortage of human capital can be the source of higher wages in Britain in the mid-eighteenth century and its abundance may be the source of later innovations (e.g. in USA). So human capital proxies were used. We have tested the enrollment ratios in primary, secondary and tertiary schools (as percentage of the total population), 5 years before the date for each innovation in each country. These data were collected from Mitchell (1998). The returns for the innovation may also be determined by the access to a larger market. Population, Openness and GDP per capita were the used proxies for the scale effect. While population and openness were collected from Mitchell (1998), GDP per capita is from the Maddison Project (Bolt and van Zanden, 2013). To account for the idea that institutional changes may have occurred earlier to allow for an increase in the innovation activity we use the overall knowledge presented in the country in 1500, a variable developed by Comin et. al. (2010). This variable is fixed on different pairs technology-year of invention, so it can be also regarded as country fixed effect which allow to account for other institutional country-specific changes. Proximity and distance have been argued to be a determinant of diffusion of technologies (e.g. Comin et. al., 2013; Spolaore and Wacziarg, 2012). The reason for this is the diffusion of ideas through the borders of the countries, which could happen even in earlier times. For instance, Mokyr (2005) studied the flow of scientists between European countries as a contribution to understand the Industrial Revolution. Communication between different scientists is easied by physical and cultural proximity. Cultural proximity can also be a determinant of similar values, norms, preferences, which shape the demand for a given good and so make more profitable to use of a certain technology. We also use physical distance (from Mayer and Zignago, 2011) and genetic distance (from Spolaore and Wacziarg, 2009) to UK and to USA as explanatory variables to innovations. Table 4.1 summarizes variables and sources.

Table 4.2 presents descriptive statistics for the main variables.

4.3 Results

We begin by considering a probit regression for the dummy of the innovator with the proxies we selected for human capital (primary education enrollment five years before the year of innovation), for scale (total population measured five years before the year of innovation), for flow of ideas, international spillovers (openness ratio measured in the year of innovation) and a proxy for the accumulated previous knowledge (which can also account for previous political and institutional reforms) (the average of the technological adoption index in 1500). For the human capital proxies, we have also tested school enrollments in secondary schools (E2 in Table 4.1) and enrollment ratio in colleges and universities (E3 in Table 4.1). However these two variables proved to be always statistically insignificant in regressions, so we have dismissed them from the results presentation. For the scale effect we also tested GDP *per capita*. Similarly, GDP *per capita* proved to be non-significant in these regressions. Table 4.3 present this benchmark regression and also the marginal effects of each variable in determining if a country was innovator or not. Column (1) shows a strong and positive scale effect related to the population size, meaning that a 2.7 increase in the number of persons will increase the probability of innovate in 4.1%. However the effect of education is small and negative. An increase in 1% in the enrollment ratio implies a decrease in the probability to innovate in 0.113%. This result is consistent to the ones that argue that the first escape from the Malthusian

Table 4.1: Variables

Dependent Variables	Name	Measure (source)
Dummy for Innovator	DI	0 or 1. Various Sources (see Appendix A).
Age of Innovation	AI	2000-Year of Innovation. Various Sources (see Appendix A).
Explanatory Variables	**Name**	**Measure (years and source)**
Education (Primary)	E1	Enrollment in Primary Schools/Population five years before the year of innovation. Mitchell (1998).
Education (Secondary)	E2	Enrollment in Secondary Schools/Population five years before the year of innovation. Mitchell (1998).
Education (Tertiary)	E3	Enrollment in Colleges and Universities/Population five years before the year of innovation. Mitchell (1998).
Openness	open	(Exports+Imports)/GDP in the year of innovation. Mitchell (1998).
Population	Pop	Logarithm of the total population five years before the year of innovation. Mitchell (1998).
GDP per capita	GDPpc	Logarithm of the GDP per capita in the year of innovation. Maddison Project (Bolt and van Zanden, 2013).
Technology in 1500	tr3	Average of the sectoral technology adoption indexes in period in 1500 AD. Comin et. al. (2010).
Distance to UK	Dist UK	Distance to United Kingdom (distance between capitals) in kilometers. Mayer and Zignago (2011)
Distance to USA	Dist USA	Distance to United States of America (distance between capitals) in kilometers. Mayer and Zignago (2011)
Genetic Distance to UK	GenD UK	Weighted FST genetic distance between United Kingdom and another country. Spolaore and Wacziarg (2009)
Genetic Distance to USA	GenD USA	Weighted FST genetic distance between United States of America and another country. Spolaore and Wacziarg (2009)

Table 4.2: Descriptive Statistics

Variable	Mean	Std. Dev.	Min	Max
(a) Dummy for Innovator	0.7692	0.26659	0	1
(b) Year of Innovation	1894.714	77.01409	1605	1994
(1) Primary Education	0.37377	0.33662	0.11115	0.99835
(2) Secondary Education	0.37223	0.22453	0.1112	0.99388
(3) Tertiary Education	0.39025	0.22593	0.1	0.99619
(4) Openness	24.12396	78.05917	0.15153	935.8247
(5) Population	64900000	108000000	438000	822000000
(6) GDP per capita	5309.041	4707.123	474.9642	24312.79
(7) Technology in 1500	0.74557	0.29821	0	1
(8) Distance to UK	3941.77	4695.531	0	17001.95
(9) Distance to USA	7931.33	3614.078	0	15961.95
(10) Genetic Distance to UK	235.9719	324.759	0	1239.689
(11) Genetic Distance to USA	441.9588	292.2913	0	1299.367

trap was due to scale and not to human capital.[1] The technological knowledge existing in the country in 1500 shows a puzzling negative effect with a marginal effect of less 1.53% probability of innovate for each increase of 0.1 in technology (the variable oscillates between 0 and 1). Column (2) repeats the regression but substituting population for GDP per capita. In this case, the only important change is that GDP per capita becomes statistically non-significant and openness becomes statistically significant with a negative sign with a very small marginal effect, meaning that openness to trade have a small negative effect on innovations and population is in fact the best proxy for the scale effect. We also consider subsets of the complete sample in regressions that we present in columns (3), (4) and (5). In particular, due to their leadership role in the first and the second industrial revolution, we exclude UK and USA from the sample and analyze if results are maintained or not. All the regressions confirm the high importance of the scale effect due to population, a small negative effect of education and a negligible effect of openness. The negative and significant effect of technology in 1500 is reverted to a significantly positive effect when USA is dropped from the sample. This is explained because the USA is the only great innovator in the post-industrial revolution that had not a significant endowment of knowledge in 1500. Thus, when the USA is dropped from the sample, a strong positive effect of technology in 1500 appears, meaning that more 0.1 in the technological index would increase in nearly 3% the probability to innovate. On the contrary, the presence of the UK increases the importance of technology in 1500 as a explanatory variable for technologies after the sixteenth century. Thus the exclusion of both results on a non-significant result.

As a second step, we have included variables linked with the notion of distance (physical and genetic) in the explanation of innovations (see Table 4.4). The analysis of these extended regressions allow for the following conclusions. First, the introduction of variables for the distances to UK and USA decrease the statistical significance of education and openness and and slightly decreases the quantitative significance of the scale effect. Now, on column 2, 2.7 more persons in the country increase the probability to innovate in only 1.8% (compared with nearly 4.1% obtained earlier). GDP per capita continues to be non-significantly related to innovations, as before. The main change however is the significance of technological knowledge in 1500. This variable is now positively and significantly related to innovation and with a huge quantitative effect. In fact, a change in 0.1 in the index of technology would imply a 3.8% to 6.2% increase in the probability of innovating. This result supports the argument according to which countries innovated in the last centuries because they have previously built knowledge and institutions that allowed innovations to be paid for. The introduction of distances to UK and to USA present also interesting results. Paradoxically, the distance to UK and proximity to the USA enhances the capacity to innovate, ceteris paribus the effects of other regressors. In fact, more 100km distance to the UK imply 0.2%-0.3% more probability of innovate. On the contrary, more 100km distance to the USA decreases the probability of innovation between 0.4% to 0.5%. Even the genetic distance between populations of the innovative countries and UK and USA seem to have even bigger effects. More 100 in genetic distance to the UK (roughly 1/3 of the genetic distance between the USA and the UK) imply more 7%-9% more probability of innovating.

One may wonder if these results are due to the presence of observations from UK and USA in the sample. So we tried to run the same regressions eliminating USA, eliminating UK and eliminating both. When we eliminate USA from the sample, physical distances to USA and UK remain highly significant and the positive and significant signs of population and technology in

[1] Easterlin's (1981) data show that a noticeable increase in British primary education occurred only until the second half of the 19th century, almost a century after the onset of the Industrial Revolution.

Table 4.3: Benchmark Probit Regressions

	(1)	(2)	(3)	(4)	(5)
	Scale Effect: Population	Scale Effect: GDP p.c.	without UK	without USA	without UK and USA
Prim. Education	-0.852**	-0.624*	-1.974**	-0.678*	-1.425
	(0.403)	(0.322)	(0.929)	(0.388)	(0.927)
	[-0.113**]	[-0.084**]	[-0.134***]	[-0.029]	[-0.026]
Openness	-0.001	-0.019**	-0.001	-0.003	-0.0014*
	(0.001)	(0.009)	(0.001)	(0.006)	(0.0008)
	[-0.000]	[-0.002**]	[-0.000]	[-0.000]	[0.001]
Population	0.309***	-	0.500***	0.441***	0.646***
	(0.070)		(0.114)	(0.089)	(0.159)
	[0.041***]		[0.034***]	[0.019***]	[0.012*]
GDP p.c.	-	0.114	-	-	-
		(0.106)			
		[0.015]			
Technology (1500)	-1.152***	-1.356***	-1.333***	6.922***	6.232
	(0.288)	(0.301)	(0.324)	(1.857)	(3.932)
	[-0.153***]	[-0.182***]	[-0.091*]	[0.299***]	[0.115]
Pseudo R^2	0.220	0.225	0.354	0.133	0.1753
Pseudo Log-Likelihood	-103.5	-123.07	-72.79	-58.50	-35.99
Number Obs.	385	432	339	340	294

Note: Level of significance: *** for p-value<0.01; **for p-value<0.05;* for p-value<0.1. Values between parentheses are standard errors.
Values in squared brackets are the marginal effects. All Models include a constant.

1500 are maintained (only marginally for technology in 1500). Moreover, the genetic distance statistical significance disappears. When we eliminate UK from the sample, physical distance to UK increases the probability to innovate and physical distance to USA decreases the probability to innovate, a result that is almost replicated when genetic distance is considered. Population and technology in 1500 remain with a statistically significant (positive) effect (in the case of population it is only marginally significant in regressions that consider genetic distance). When we eliminate both UK and USA, physical distance to UK and proximity to the USA increases the probability to innovate of other countries, recovering the initial effect described in Table 4.4. On the contrary, genetic distance turns out to be non-significant.

Overall, when the effect of distances (both geographical and genetic) to the innovation leaders (UK and USA) is taken into account, positive significant effects of scale (population) and previous accumulated knowledge are evident on the probability of innovate. Significant positive effects of distance to UK and proximity to the USA have also been uncovered. However, eliminating the effect of considering these countries in the sample, only the geographical distance (and not the genetic one) seems to influence the probability of innovate.

4.3.1 Determinants of Earlier Innovations

In this Section we present results for OLS regressions with a dependent variable that intends to measure how early an innovation occurs. With this, our aim is to test the determinants of some innovations being invented earlier than others. The dependent variable is 2000-year of invention. We tested the same explanatory variables as before. However, due to the lack of degrees of freedom, we successively eliminated variables that proved to be non-significant.

Table 4.5 shows the selected regressions. Column (1) confirms a small negative effect of primary

Table 4.4: Extended Probit Regressions

	(1)	(2)	(3)	(4)
Primary Education	-0.491	-0.392	-0.483	-0.274
	(0.373)	(0.377)	(0.345)	(0.377)
	[-0.032]	[-0.019]	[-0.043]	[-0.020]
Openness	-0.004	-0.006	-0.0117	-0.013
	(0.006)	(0.006)	(0.008)	(0.009)
	[-0.0003]	[-0.0003]	[-0.001*]	[-0.001]
Population	0.336**	0.372***	-	-
	(0.164)	(0.124)		
	[0.022]	[0.018***]		
GDP per capita	-	-	0.005	-0.083
			(0.145)	(0.141)
			[0.00046]	[-0.0062]
Technology in 1500	5.953**	11.708***	4.725**	8.309***
	(2.641)	(1.604)	(1.983)	(1.766)
	[0.382***]	[0.563***]	[0.424***]	[0.620***]
Physical Distance to UK	0.0004**	-	0.0003***	-
	(0.0002)		(0.0001)	
	[.00002***]		[.00003***]	
Physical Distance to USA	-0.0006***	-	-0.0006***	-
	(0.0002)		(0.0002)	
	[-.00004***]		[-.00005***]	
Genetic Distance to UK	-	0.0152***	-	0.0120***
		(0.002)		(0.002)
		[0.0007***]		[0.0009***]
Genetic Distance to USA	-	-0.0178***	-	-0.014***
		(0.002)		(0.003)
		[-0.001***]		[-0.001***]
Pseudo R^2	0.336	0.346	0.332	0.358
Pseudo Log-Likelihood	-88.08	-79.11	-106.06	-93.1
Number Obs.	385	356	432	401

Note: Level of significance: *** for p-value<0.01; **for p-value<0.05;* for p-value<0.1. Values between parentheses are standard errors.
Values in squared brackets are the marginal effects.

Table 4.5: OLS Regressions for Earlier Innovations

	(1)	(2)	(3)	(4)
Primary Education	-182.82**	-201.04	-	-
	(74.34)	(137.84)		
Openness	-0.624	-	-	-
	(0.797)			
Population	-73.896***	-59.475***	-47.948***	-51.406***
	(3.35)	(12.215)	(6.854)	(8.033)
Technology in 1500	-81.214***	-82.33***	168.833***	140.455***
	(22.023)	(16.632)	(48.075)	(27.482)
Physical Distance to UK	-	-	0.0115***	-
			(0.004)	
Physical Distance to USA	-	-	-0.0158***	-
			(0.003)	
Genetic Distance to UK	-	-	-	0.2199***
				(0.035)
Genetic Distance to USA	-	-	-	-0.2489***
				(0.039)
R^2	0.90	0.705	0.608	0.642
F-stat	157.24	10.68	29.07	15.18
Number Obs.	42	43	61	49

Note: Level of significance: *** for p-value<0.01; **for p-value<0.05;* for p-value<0.1. Values between parentheses are standard errors.

education to innovate earlier (other levels of education were tested, as before, but they proved to be always non-significant), a non-significant effect of openness and a puzzling negative effect of technology in 1500. There is also a negative scale effect. This must be because when early innovations appeared, population levels were smaller than when latter innovations appeared, a natural effect due to the fact that population was growing in most of the analyzed periods. When openness - the non-significant effect - is dropped (column 2) education also becomes statistically non-significant and the other effects remain qualitatively similar to those in column (1). Columns (3) and (4) dropped education and included physical distance and genetic distance, respectively. An increase in 2.7 in population would have deterred innovations by 47 to 51 years, a strong effect. However, the sign of the ancient technology effect is switched to a more intuitive effect. Now, more technology in 1500, say more 0.1, fostered innovations in 14 to 17 years. The introduction of distance has again interesting effects. In fact, distance to UK and proximity to USA (both physical and genetic) did foster innovations. While the quantitative effect of physical distance is rather modest (a 100 km distance from the UK have fostered innovations in 1 year while 100 km closer to the USA fostered innovations in 1 year and 1/2), the quantitative effect of genetic distance is more important. In fact, 100 genetic distance from the UK would have deterred innovations for more 22 years while 100 genetic proximity to the USA would have fostered innovation in 25 years.

4.4 Conclusion

We collected the invention dates for more than 100 inventions from the seventeenth century around the world. With that we studied the determinants of the probability of the innovations occur in a given country, trying to contribute to the literature that explains the triggers of the industrial revolutions.

We found evidence according to which the scale of the country, measured by its population,

is an important determinant of the probability to innovate. This strong effect means that nearly more 30 inhabitants in a country could increase the probability to innovate from 12% to 41%. Our results show a small negative effect of education, reflecting the relatively lower importance of education as a source of innovations. This is according to the opinion of some economic historians which argue that the advent of formal education was posterior to the rise of innovations during the industrial revolution. Openness is rarely a significant determinant of inventions and when it is, it appears with a negative sign. Innovations from 1600 onwards are also related to previous technological knowledge in the country. We found a negative effect if the USA is included in the sample and a positive effect otherwise. This reflects the fact that the United States developed a strong industrial revolution without any substantial technological development in 1500. A standard-deviation increase in technology in 1500 (0.3) would increase the probability to innovate from 3% to 9% (when a positive effect occurs).
We also found evidence for the influence of geographic distance and genetic distance on the probability of inventions. Generally and interestingly, distance to the UK and proximity to the USA increased the historical probability of innovation for a given country. In particular, geographical distance statistical significance is robust to all specification changes we have performed. However, the quantitative effects are small. An increased distance of 100 kilometers to the UK imply 0.2%-0.3% more probability of innovate. On the contrary, more 100km distance to the USA decreases the probability of innovation between 0.4% to 0.5%. More 100 in genetic distance to the UK (roughly 1/3 of the genetic distance between the USA and the UK) imply more 7%-9% more probability of innovating.
Finally, we have tested the influence of the same regressors on a variable that intends to measure how early innovations occurred. Now, a rise in population by 30 persons would have deterred innovations by 47 to 51 years. More technology in 1500, say more 0.1, foster innovations in 14 to 17 years. Distance to UK and proximity to USA (both physical and genetic) did foster innovations. While the quantitative effect of physical distance is rather modest (a 100 km distance from the UK foster innovations in 1 year while 100 km closer to the USA foster innovations in 1 year and 1/2), the quantitative effect of genetic distance is more important: 100 genetic distance from the UK would have deterred innovations for more 22 years while 100 genetic proximity to the USA would have fostered innovation in 25 years.
This article presents evidence that tends to favor explanations of earlier innovations linked with scale and previous technological endowment but tend to dismiss explanations linked with formal schooling.

Chapter 5

Income Inequality, TFP, and Human Capital

5.1 Introduction

Understanding the causes of inequality is fundamental to indicate possible policy measures that ensure that the increased production and income of societies can be better shared among the whole population. Reducing inequality is important not just to achieve a fairer distribution of income and address the social concerns that widening disparities in income raise, but also to ensure a good environment for growth. As has been seen in some countries, these social concerns can lead to social instability. Income inequality may itself limit the growth potential of economies as social, economic, and political instability caused by inequality is associated with slower growth. Even in democracies, an increase in inequality may contribute to elect politicians that are against openness and globalization, which may deter the world integration process which is known to have positive effect on the growth prospects of the economy.
There is a fruitful theoretical literature interested in explaining the rise of inequality in the second-half of the twentieth century (mainly in the USA) together with the rise in the supply of human capital. Skill biased technical change and capital-skill complementarity have been crucial to explain this phenomenon. Generally, according to this theory, skill-premia increase due to two effects. First, the skill premium would reflect the productivity difference between sectors. Second, with full capital mobility, factor price equalization requires capital to flow to the sector operating the new technology, and thus workers in the new technologies sectors are endowed with more capital, which boosts their relative wages (Acemoglu, 2002a, 2002b, 2003). An alternative development has argued that the diffusion of IT - General Purpose Technologies - may have raised the demand for adaptable skilled workers and made vintages of capital more adaptable. Therefore, this increases the premium of workers that show a lower learning cost and can adapt quickly from one sector to another. These ideas have been formalized by Galor and Tsidon (1997), Greenwood and Yorukoglu (1997), Caselli (1999), Galor and Moav (2000) and Aghion, Howitt, and Violante (2002). Theoretically, skill-biased technological change is explained by the proportion of skills (education) in the economy, and wage inequality (typically measured by the wage ratio between skilled and unskilled workers) is proportional to the proportion of skills in the economy. Education is thus seen in the theory as a determinant of more technical change (and consequently growth) and more inequality.
Whatever the explanation is for the rise in inequality and its relationship to technology and human capital, there is little quantitative literature on the issue, as pointed out by Hornstein, Krusel and Violante (2005:1361). In fact, empirical attempts to evaluate the relationship are mostly country-specific as, e.g. Ding et al. (2011) and Rattsø and Stokke (2013) dealing with the effect of technology, and in Birchenall (2001) dealing with the effect of human capital. Micro evidence on the relationship between education and income inequality is mixed. While Martins and Pereira (2004) found a positivesign for the effect of education returns in inequality due to an increase in returns to education throughout the wage distribution for 16 European Countries, Wang (2011) found returns to education in China that are more pronounced for individuals in the lower tail of the earnings distribution than for those in the upper tail, in stark contrast to

the results found in some developed countries.
We have only found two papers that evaluated this relationship using a large cross-section of countries. Barro (2000) presents fixed-effects estimations of equations of the Gini index on covariates such as GDP and GDP squared, schooling, democracy index, openness, rule of law index and several dummies. In his fixed-effects estimations, dummies for income or spending and secondary schooling are negatively related to inequality and higher schooling and openness are positively related to inequality (with significant coefficients). Primary schooling and the dummy for individual or household data are insignificantly related to the Gini coefficient. There is a strong inverted-U relationship with GDP (the so-called Kuznets curve) in Barro's estimations. Recently, Jaumotte, Lall, and Papageorgiou (2013) re-assessed the determinants of inequality. They focus on the effect of globalization on inequality but avoid the relationship between inequality and GDP. They conclude that trade globalization decreases inequality while financial globalization increase inequality. Moreover, information and communication technologies and credit deepening increases inequality while the share of industry in the economy decreases inequality. Interestingly, education variables and initial GDP (when included) are insignificantly related to inequality.
As can be noted, empirical evidence coming from a large cross-section of countries has quite ambiguous results regarding the determinants of inequality and does not confirm theories in crucial aspects such as the influence of education and technology. However, much criticism has affected data on inequality around the world. In fact, greater coverage across countries and over time is available from these sources only at the cost of significantly reduced comparability across observations. There are currently three different projects that collect and make publicly available inequality data for many countries and periods around the world: the Luxembourg Income Study (LIS), the dataset assembled by Deininger and Squire (1996) for the World Bank (WIID) - recently updated and upgraded by the WIDER (World Institute for Development Economic Analysis) project, and the most recent standardized World Income Inequality dataset (SWIID), by Solt (2009). The LIS, which was used by Jaumotte, Lall, and Papageorgiou (2013), has generated the most-comparable income inequality statistics currently available but covers relatively few countries and years. The Deininger and Squire dataset and its successors, used by Barro (2000), on the other hand, provide much more observations, but only at a substantial loss of comparability. Solt (2009) implemented a sequence of steps in order to standardize income inequality data and provide data with more ample coverage than the WIID but at the highest quality as in LIS. However, in the process of standardization, not all countries had the sufficient data in the original sources. To handle this, Solt (2009) also calculated a standard-error of each Gini coefficient to account for the remaining uncertainty in data.
This paper contributes to our knowledge of the relationship between human capital, technology and inequality in two crucial ways: first, it uses a large database on inequality, based on the Standardized World Income Inequality dataset, and combines it with the most recent data for human capital and TFP to explain cross-country patterns of inequality; second, for the first time, it takes into account country heterogeneity, cross-country dependence, and endogeneity to common factors in evaluating the effects of human capital and TFP on inequality. The exploration of a large dataset of over 150 countries across more than 50 years (since 1960) allowed us to explore issues such as panel heterogeneity, cross-country dependence and time-series features, such as stationarity and causality, which are absent from earlier contributions. Exploring the heterogeneity of data concerning the determinants of inequality is especially important since the effects of different inequality determinants may differ considerably from country to country. In fact, and to give a few examples, the effect of technology adoption may differ if

the country is on the technological frontier or lagging behind; the effect of human capital may differ between countries where brain-drain is more evident than in others; and the effect of openness may depend crucially on the level of integration and on the market size of the country. In general, historical and institutional (e.g. labor market related) country-specific factors that are not simply captured by fixed-effects estimations, are in fact dealt through heterogeneous panel estimations.

Our main conclusions point out to a clearly significant, worldwide relevant, positive effect of human capital on inequality, an effect that is stronger for the developed world. On the contrary, our results indicate that the effects of technology and openness may be quite different from country to country, as well as dependent on different specifications. Overall, the common factors framework dismiss the existence of a Kuznets curve.

The remainder of the paper is organized as follows. Next, in Section 5.2 we describe our dataset. In Section 5.3 we describe our estimation strategy. In Section 5.4 we present our results, beginning with detailed evidence for cross-country dependence, stationarity, and causality and then showing the results from several different specifications based on heterogeneous panels methods. Section 5.5 concludes.

5.2 Sources and Data

We use data from the Standardized World Income Inequality database (SWIID), version 4.0, from Solt (2009), for the Gini coefficient.[1,2] These include data on the Gini coefficient using post-taxes and post-transfers income (the net definition) and on the Gini coefficient using pre-taxes and pre-transfers income (the market definition), and the respective standard-errors by country and year. Previous data on inequality have presented variables divided by the type of underlying measure of inequality (income or consumption) and by the quality of data (e.g. defining different quality levels). Solt (2009) maintained the same concerns within their dataset. They divide data in net and market Gini indexes which may be roughly matched with consumption and (net) income Gini indexes, by one side, and (gross) income, by the other Gini indexes. Additionally they provide the data with a standard-deviation, which intends to measure uncertainty in data, basically due to less availability of underlying data to calculate inequality measures in some countries. Thus, in his database, this can be interpreted as the information about the quality of the underlying data. In the majority of the analysis made in the paper, we will use a quality-adjusted measure of the SWIID gini coefficient which is simply given by dividing the Gini coefficient by the respective standard-deviation, provided by Solt (2009).[3]

[1] Available at $http://thedata.harvard.edu/dvn/dv/fsolt/faces/study/StudyPage.xhtml?studyId=36908$. This is the first time this source for inequality data is used to access the relative importance of the determinants of inequality. We explained above the reasons why this choice is superior to the previously used data.

[2] In a working-paper version of this article, we compare some results with inequality data coming from the Word Income Inequality database (WIID2c). In doing so, we followed some strict criteria to select data, separating Gini coefficients from net income, consumption and gross income and preferring data with wide coverage and higher quality. In that analysis, we also made clear that SWIID have more than four times the number of observations than the measures coming from the WIID, making SWIID more suitable (if not the unique suitable) for being studied with heterogeneous panel data methods.

[3] The uncertainty-corrected measure is $\frac{GINI}{sd(GINI)}$, where *GINI* is the Gini index provided by SWIID and *sd(GINI)* is the standard-deviation of the Gini index, also provided by the SWIID and that corrects for uncertainty or measurement error within the sources. Later on, on the Discussion section, we discuss the results obtained with an alternative uncertainty-corrected measure.

Table 5.1: Descriptive Statistics

Variable	N. Obs	Mean	Std. Dev.	Min	Max
Gini (net)	4597	3.5923	0.2960	2.7324	7.3871
Gini (market)	4597	3.7395	0.2234	2.8367	4.3740
Gini (net) - value/sd	4597	3.5613	0.9786	1.2658	9.5894
Gini (market) - value/sd	4597	3.2479	1.0049	1.0747	9.5410
Human Capital	6797	0.6905	0.3160	0.0198	1.2861
TFP	4994	0.5254	0.5287	-3.5389	1.1222
Openness	7760	1.1645	1.1020	-12.7415	3.2061
GDP per capita	7760	8.2779	1.1891	4.8890	10.9961

Notes: Gini variables are from SWIID - Standardized World Income Inequality Database, from Solt (2009). Human Capital, TFP, Openness = (Exports+Imports)/GDP - and GDP per capita are from PWT 8.0. When value/sd is indicated it means that the Gini coefficient is divided by its standard-error, a measure to account for uncertainty in the data for each country-year pair. All variables are in natural logarithms.

We use GDP *per capita*, openness, human capital index, and TFP index from Penn World Tables (PWT), version 8.0 (Feenstra et al., 2013).[4] Human capital in PWT 8.0 is measured by a 'Mincerian' combination of years of schooling (from Barro and Lee, 2013, version 1.3) and returns to education. The results from Psacharopoulos (1994) show that returns from schooling decrease across years of schooling. As the influence of human capital in inequality arguably changes through years of schooling (Barro's results show negative signs for primary and secondary schooling and positive signs for tertiary schooling) and returns from schooling are essential to understand income inequality, we think this variable is the most appropriate human capital measure to enter in inequality regressions. In fact, as human capital measures corrected for returns for education weights more lower levels of education, they correct underestimations of human capital in less developed countries. Lower levels of education in less developed countries may have more influence in decreasing wage inequality than they have in more developed countries. The human capital measure provided by the PWT 8.0 is the one with the highest coverage until now, as it not only corrects years of schooling by different returns by levels of education, but it is also interpolated to provide annual measures. It is worth noting that returns to education differ between levels of education but not between different countries or years as these alternatives would result in lower coverage.

TFP is available in PWT 8.0 both as a ratio to the USA=1 level and on constant national prices. We construct our index departing from a final TFP level (related to the USA) in 2011 and then deflating year by year using growth rates of the national currency measure of TFP. This allows us to have a PPP measure of TFP that is independent of the USA level (at an year-by-year basis) in the time-series analyzed.[5] Contrary to Barro (2000) but similar with Jaumotte, Lall, and Papageorgiou (2013), we used annual data.

We end up with an unbalanced panel database of 156 countries with an average of 31 years per country, from 1960 to 2011.[6] Table 5.1 shows descriptive statistics for the variables included in the analysis.

5.3 Estimation and Methods

The first issue to deal with the estimation is to choose the explanatory variables to the equation for inequality. The theory explains inequality through skill-biased technical change and

[4] Available at $http://www.rug.nl/research/ggdc/data/penn-world-table$.

[5] We began with the year 2011 in order to maximize the available data for the TFP index.

[6] 31 observations per country is the average number of time-series per country considering the pool of the mentioned variables although some variables may include nearly 50 years per country.

thus human capital and technology seem to be the main theoretical determinants of inequality. Additionally, openness to trade in the theory increases inequality, also suggesting that openness ratio may be considered also as a determinant of inequality. Thus, theory points out three main determinants of inequality: human capital, technology and openness (Acemoglu, 2002a,b). One must note however that according to the theory, technology is endogenous as the direction of technical change is also determined by human capital. From the observation of previous empirical contributions from Barro (2000) and Jaumotte, Lall, and Papageorgiou (2013) one may retain that common regressors should be linked with technology, human capital and openness. While Barro (2000) also include the estimation of the Kuznets' curve, rule of law and democracy indexes and several dummies, Jaumotte, Lall, and Papageorgiou (2013) includes several variables for trade and financial globalization, shares and productivity series for industry and agriculture and private credit. Chakrabarti (2000) studied the effect of openness to trade and inequality but do not consider the effects of human capital and technology explicitly. We choose to estimate a more parsimonious specification.[7] Our estimation method hereinafter is the common factor framework for heterogeneous panels from Pesaran (2006) and followers. Our baseline specification is thus as follows:

$$gini_{it} = \beta_{1i} hcap_{it} + \beta_{2i} TFP_{it} + \beta_{3i} Open_{it} + \lambda_i' \mathbf{f}_t + \alpha_i + u_{it} \quad (5.1)$$

where $gini$ is the natural logarithm of the Gini coefficient, TFP is the natural logarithm of a measure of total factor productivity, $hcap$ is the the natural logarithm of the human capital variable, $Open$ is the the natural logarithm of the openness ratio, α_i is the country fixed-effect, $\mathbf{f}_t$ is the vector of unobservable common factors, λ_i' is the associated vector of factor loadings and u_{it} is the error term. As can be observed from (5.1), each coefficient is country-specific, thus allowing for complete heterogeneity in the estimation. Additionally, as each regressor can also depend on the common factor, the method is also robust to endogeneity of the observable factors toward the common factors determining inequality. As Pesaran and Tosetti (2011) explain, this method is robust to non-stationarity in both observables and non-observables and works well in the presence of weak and/or strong cross-sectionally correlated errors.[8] As the analysis in Jaumotte, Lall and Papageorgiou (2013) might indicate, we suspect that the Gini coefficients, financial openness, and technological development may well be non-stationary and heterogeneous among different countries. Finally, we may consider that technology adoption is being determined by the same phenomena as inequality, say by common factors such as globalization or the entry of China into the world market, technology thus being an endogenous variable.[9] These are the reasons why we will apply the Pesaran (2006) estimator for heterogeneous panels.

[7]We performed specification testing against the existence of the Kuznet's curve (GDP *per capita* and GDP *per capita* squared) and our results indicate that those variables are not significant when added to our benchmark specification. Additionally, the inclusion of GDP *per capita* as a explanatory variable for inequality would imply obvious multi-collinearity with other variables, such as human capital and TFP. These results are available upon request.

[8]There are not many empirical applications with those heterogeneous panel methods. Notable exceptions are the recent papers from Markus Eberhardt and co-authors (Eberhardt and Presbitero, 2014; Eberhardt and Teal, 2013a, 2013b and Eberhardt, Helmers, and Strauss, 2013). Eberhardt and Teal (2011) explain why the standard cross-country regression framework and its panel cousins needs to be reconsidered. None of these papers deal with income inequality.

[9]For complete arguments toward reconsideration of traditional econometric methods to study moderate-T dimensional panel data of countries, see Eberhardt and Teal (2011).

Table 5.2: Cross-sectional dependence test

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variable	Gini Net Income (>30)	Gini market (>30)	Gini Net Income (>30, ./sd)	Gini Market Income (>30, ./sd)	Human Capital	TFP	Openness
$CD\ Test$	23.33***	19.79***	96.40***	79.47***	554.05***	53.81***	240.32***
p-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Number of Countries	82	82	82	82	128	106	155

Note: >30 indicates that only cross-sections with more that 30 time-series observations are included. Level of significance: *** for p-value<0.01; **for p-value<0.05;* for p-value<0.1. ./sd indicates when the Gini coefficient is divided by the source standard-deviation to account for data uncertainty. All variables are in natural logarithms.

5.4 Empirical Results

Our results section begins by presenting evidence of the time-series properties of inequality. Due to unbalance and holes in several time series, to perform some of those tests, we limit our variable of interest such that we include only countries with more than a given number of time-series observations (30) in the Gini index series.[10] We consider both the Gini coefficient as provided by the source as well as an uncertainty-corrected version of the Gini coefficient which consists of dividing the coefficient by the standard-deviation (also provided by the source).

These new data on inequality provide, for the first time, the means for analyzing time-series features in a reasonable set of countries. This analysis occupies Sub-Sections 5.4.1 and 5.4.2.[11] Then, in Section 5.4.3 we present evidence on the relationship between human capital, TFP and openness in inequality in a heterogeneous panel setup and several robustness analyses. Section 5.4.4. discusses the results.

5.4.1 Initial Analysis: cross-country dependence and stationarity

The standard literature on the panel data analysis assumes cross-sectional independence. However, there are several reasons why cross-sectional dependent error structure can arise in a large panel data of countries. Such cross-correlations can arise due to omitted common factors that affect the evolution of inequality, including technological cross-country spillovers, migration of high and low skilled workers and integration in international markets. As Pesaran and Tosetti (2011) write, "conditioning on variables specific to the cross-section units alone does not deliver cross-section error independence, an assumption required by the standard literature on panel data models", the one that has been applied in the existing analyses of the determinants of inequality. Table 5.2 shows results for the cross-sectional dependence test from Pesaran (2004) which tests the null of no cross-sectional dependence.

[10]This would be the minimum number or time-series observations for the Gini index. However, due to the unbalanced nature of the panel, the time-series observations that effectively enter in regressions may be lower than 30.

[11]It should be noted however, as stressed by Eberhardt and Teal (2011), that most of the unit-root and cointegration tests have low power in panels of moderate dimension such as the one under analysis. This does not invalidate that their results constitute important motivation to choose a heterogeneous common factor approach that is indeed appropriate to deal with moderate N, moderate T panels, typical in macroeconomic analysis.

These tests constitute overwhelming evidence that the series of inequality (as well as their main determinants) are cross-country related, thus inducing bias on estimations assuming cross-country independence. It is interesting to note that the series with the highest cross-dependence test is human capital, following by openness. Also worth noting is that the uncertainty corrected measures of the Gini coefficient present higher values for the test than the original Gini coefficients, indicating an increased correlation between countries in these uncertainty-corrected measures. Although we provided results from the Gini coefficient from the market approach in this Table 5.2, from now on we will concentrate on the most interesting variable: the Gini coefficient from post-tax and post-transfers income. This variable incorporates the effects of progressive tax systems and is close to a measure of inequality related to disposable income.[12]
Another issue to be dealt with is the integration level of the series, i.e. its stationarity or non-stationarity. It is well-known that most macro time series are non-stationary even though the issue has received virtually no attention in traditional panel regression analyses (Phillips and Moon, 2000: 264). The graphic analysis in Jaumotte, Lall and Papageorgiou (2013: 277-283) is a means for observing non-stationarity of Gini coefficients and their determinants. Table 5.3 shows unit root tests. We use the Pesaran (2007) Panel Unit Root test (CIPS) whose null is that the variable is I(1). The analysis of results – with the majority of the tests on the level variables not rejecting – points out the non-stationarity of the Gini coefficients and some of their determinants, with particularly clear results for human capital. The only determinant of inequality for which the tests clearly reject non-stationarity is Openness. These results are confirmed by the tests on the differenced variables (see Table D.1.1), which clearly reject the unit root case.
This section provides clear motivation that the heterogeneous panels unobserved common factors framework from Pesaran (2006) and followers is appropriate to analyze inequality determinants. The availability of data in quality and quantity allow for its correct implementation.
The next section explores the causal relationship between inequality and human capital.

5.4.2 Initial Analysis: causality between education and inequality

Trade and productivity (or technology) as determinants of inequality have been widely studied and the causal relationship from openness and technology to inequality is well founded in theory (see e.g. Hornstein, Krusel and Violante, 2005, Chakrabarti, 2000, and Richardson, 1995). However, the causality path from human capital to inequality is not so well founded. Despite the tremendous emphasis on the role of human capital in the skill-biased technological change and general purpose technology literatures, there are some microeconomic arguments that come from the economics of education field suggesting that inequality may decrease incentives to educate and thus decrease human capital (Stocké et al, 2011 and Gutiérrez and Tanaka, 2009 are good examples that emphasize the causality channel from inequality to education). It is important then to evaluate evidence in our data from the causality channel between human capital and inequality. We do this using a cointegration test for the null of no cointegration, the Westerlund (2007) test. Table 5.4 presents the tests when the causality is evaluated between

[12]Variables linked with disposable income have also been the focus of earlier papers. Barro (2000) uses a dummy to account for differences from the net income and consumption definition and gross income definition. This dummy is highly significant indicating that these variables measure in fact different phenomena. Jaumotte, Lall and Papageorgiou (2013: 276) also express concern about jointly analyzing income and expenditure-based Gini indexes. Results obtained with the *market* Gini coefficient (and its uncertainty-corrected version), which can compare with the ones presented in the paper can be provided by the authors.

Table 5.3: Panel Unit-Root tests

Variable	Lag	(1) Gini Net Income (>30)	(2) Gini Net Income (>30, ./sd)	(3) Human Capital	(4) TFP	(5) Openness
		Pesaran (2007) Test without Trend				
Zt-stat	0	3.08	-10.29***	17.17	-3.30***	-6.58***
p-value		(0.999)	(0.000)	(1.000)	(0.000)	(0.000)
Zt-stat	1	-0.406	-7.70***	3.51	-3.37***	-5.44***
p-value		(0.342)	(0.000)	(1.000)	(0.000)	(0.000)
Zt-stat	2	-2.39***	-2.93***	3.80	-3.27***	-2.35***
p-value		(0.008)	(0.002)	(1.000)	(0.001)	(0.009)
Zt-stat	3	1.62	-2.091***	3.15	-2.51***	-1.45*
p-value		(0.948)	(0.018)	(0.999)	(0.006)	(0.073)
		Pesaran (2007) Test with Trend				
Zt-stat	0	6.17	-5.846***	14.49	0.70	-7.65***
p-value		(1.000)	(0.000)	(1.000)	(0.758)	(0.000)
Zt-stat	1	1.23	-3.451***	5.20	0.09	-5.66***
p-value		(0.109)	(0.000)	(1.000)	(0.535)	(0.000)
Zt-stat	2	-4.45***	2.685	6.04	0.28	-2.73***
p-value		(0.000)	(0.354)	(1.000)	(0.610)	(0.002)
Zt-stat	3	0.35	2.752	6.52	1.82	-1.65**
p-value		(0.635)	(0.997)	(1.000)	(0.965)	(0.049)
Number of Countries		82	82	128	106	155
N. of Observations		3224	3224	6694	4994	7760
Avr. N. of Obs.		40.5	40.5	55.4	51.5	53.9

Note: All variables are in natural logarithms. >30 indicates that only cross-sections with more that 30 time-series observations are included. ./sd indicates when the Gini coefficient is divided by the source standard-deviation to account for data uncertainty. Level of significance: *** for p-value<0.01; **for p-value<0.05;* for p-value<0.1.

Table 5.4: Cointegration tests

		(1)	(5)	(6)	(7)	(8)
	Lag	Trend	Test Gt	Test Ga	Test Pt	Test Pa
Dependent Variable			Gini Coefficient net income (>30, ./sd)			
	1	No	-2.400***	-10.22***	-9.630***	-9.588***
p-value			(0.001)	(0.004)	(0.002)	(0.000)
	1	Yes	-2.653**	-12.64	-10.79	-11.343**
p-value			(0.049)	(0.332)	(0.156)	(0.033)
	2	No	-2.353***	-8.232	-7.195	-7.261***
p-value			(0.001)	(0.174)	(0.342)	(0.000)
	2	Yes	-2.689**	-10.952	-7.660	-8.500
p-value			(0.031)	(0.768)	(0.995)	(0.630)
Dependent Variable			Human Capital			
	1	No	-1.826	-3.711	-5.713	-1.453
p-value			(0.401)	(0.998)	(0.861)	(0.998)
	1	Yes	-1.990	-7.607	-9.765	-6.089
p-value			(0.985)	(0.999)	(0.565)	(0.985)
	2	No	-1.879	-3.855	-5.448	-1.406
p-value			(0.298)	(0.998)	(0.912)	(0.999)
	2	Yes	-1.807	-7.110	-8.696	-5.479
p-value			(0.999)	(1.000)	(0.917)	(0.996)

Note: All variables are in natural logarithms. >30 indicates that only cross-sections with more that 30 time-series observations are included. All tests include a constant. ./sd indicates when the Gini coefficient is divided by the source standard-deviation to account for data uncertainty. Level of significance: *** for p-value<0.01; **for p-value<0.05;* for p-value<0.1. Rejection of H0 in Ga and Gt tests should be taken as evidence of cointegration of at least one of the cross-sectional units. Rejection of H0 in Pa and Pt tests should therefore be taken as evidence of cointegration for the panel as a whole.

human capital and the uncertainty-corrected Gini coefficient. The intuition is as follows. If the null is rejected for a test in which the dependent variable is inequality and simultaneously the null is not rejected for a test in which the dependent variable is human capital, then human capital has a (Granger) causal effect on inequality and inequality has no (Granger) causal effect on human capital. The pattern of results clearly suggests a (Granger) causal relationship from human capital to inequality and not the other way around. This is valid for both the uncertainty-corrected measure presented in Table 5.4 and for the uncorrected measure.[13] As in previous tests, we use only cross-sections that have availability of time-series data of 30 or more periods.

The next sections present results for the influence of human capital, TFP, and openness on inequality using heterogeneous panels methods.

5.4.3 Results: baseline specification

In this section we present the results for our baseline specification in equation (5.1). Results in Table 5.5 show that, for uncorrected Gini indexes, human capital, TFP and Openness are not quite significant which may mean that there is great heterogeneity concerning effects of the three determinants across countries. Human capital is significant only in the regression for the Gini coefficient - with a negative sign when the Gini coefficient is not corrected for uncertainty and for the restricted sample with longer time-series within panels - Table 5.5, column (2) - and with a positive sign when the Gini coefficient is corrected for uncertainty - Table 5.5, columns (3) and (4). In the former case, an increase in 1% in human capital would imply a decrease of 0.27% in the uncorrected Gini coefficient. In the latter, however, a 1% increase in human capital would increase the corrected Gini coefficient from 2.4% to 3.7%. Alternatively, it can be

[13]Results for the uncorrected measure are in Table D.1.2.

said that for the same level of precision of the Gini coefficient, a 1% increase in human capital would increase the corrected Gini coefficient in values ranging from 2.4% to 3.7% (depending on the database being unrestricted or restricted to countries with time-series longer than 30 observations).
The heterogeneity of effects are indeed high especially in what the effects of TFP and Openness are concerned, as can be observed by the count of significant effects by country, provided in the Table.[14] The number of countries with significant results for each variable are usually more than 50% of the number of countries included in the regressions. The number of countries with significantly positive coefficients and the number of countries with significantly negative coefficients for TFP and openness are relatively balanced, thus yielding in general non-significant averaged coefficients.
For human capital coefficients, in the uncorrected Gini regressions, the number of countries with significant negative coefficients (39 and 27 respectively for columns (1) and (2)) are far more than the countries with positive and significant coefficients (19 and 7 respectively). For human capital coefficients, in the corrected Gini regressions, the picture is now switched: the number of countries with significant positive coefficients (43 and 35 respectively for columns (3) and (4)) are far more than the countries with negative and significant coefficients (9 and 3 respectively). The overall significance of regressors is higher in regressions with the corrected Gini index, as shown by higher significance of the Wald tests, which will be an additional reason for considering this corrected measure as our preferred one. It is also worth noting that the cross-independence of the residuals is not rejected in column (2) for the uncorrected measure. However, the value of the test indicates that cross-dependence is now much lower than the level affecting the regressors, when compared to values in Table 5.2. Additionally, the cross-independence of residuals cannot be rejected in the regression of column (4), for the corrected Gini index. It should be noted that correcting for uncertainty in the information set used to construct inequality measures has been essential to uncover a positive effect of human capital on inequality which is consistent with the theoretical literature on the issue.
It is possible now to present an idea of the countries for which significant effects were detected.[15]. In regressions presented in columns (3) and (4), the great majority of countries presents a statistically significant positive coefficient for human capital. The exceptions are Bulgaria, Burundi, Central African Republic, Cyprus, Latvia, Mongolia, Namibia, Romania, and Swazilandia. A complete matching with relatively poor countries is not possible although some of the richest countries in the world present a significantly positive effect of human capital on inequality. Those are Australia, Canada, Finland, Hong Kong, Italy, Netherlands and Norway, just to mention some of them. Countries in which TFP tends to increase inequality are, among others, France, Germany, Japan, Botswana, Bulgaria, and Chile and those with a negative effect are, for example, Austria, Belgium, Croatia, Iran, Israel, Korea, Russia, and the USA. Openness tends to increase inequality in Austria, Canada, Colombia, Estonia, Greece, Indonesia, Japan, Russia, Phillipines and the USA and to decrease inequality in Bulgaria, Jordan, Portugal, Romania, Slovenia and Taiwan. Also in these cases, it is not possible to present an *a priori* association between those countries and the respective level of income or some other common feature that may characterize them. Next, we split our sample into rich and poor countries and systematically evaluate the effects of human capital, TFP, and openness in each of the sub-samples. We used the sample median for real GDP *per capita* as the threshold to split the sample. Countries with an average of GDP *per capita* above the median would be classified

[14]We follow Eberhardt and Presbitero (2014) in showing counts of significant effects.
[15]The lists of those countries for regressions in columns (3) and (4) are listed in Appendix D.2

Table 5.5: Inequality, Human Capital, TFP, and Openness

	(1)	(2)	(3)	(4)
Dependent Variable: Gini Measure	Gini Net post-tax; post-transfer	Gini Net post-tax; post-transfer >30	Gini Net post-tax; post-transfer /.sd	Gini Net post-tax; post-transfer /.sd >30
$hcap$	-0.204	-0.272**	2.406***	3.737***
	(0.195)	(0.050)	(0.001)	(0.000)
TFP	0.001	-0.038	-0.116	-0.230
	(0.965)	(0.314)	(0.391)	(0.196)
$Open$	0.011	0.009	0.002	-0.009
	(0.431)	(0.600)	(0.963)	(0.865)
N Observ.	3300	2593	3300	2593
Avr. N Obs.	32	38.1	32	38.1
Min-Max	7-52	21-52	7-52	21-52
Number Countries	103	68	103	68
Wald	2.31	5.13	11.04**	21.64***
CD-test (res)	-	1.95* (0.052)	-	-0.28 (0.782)
Stat-test (res)	-	rejects I(1)	-	reject I(1)
sig. signs /countries for $hcap$	↗(19)↘(39)	↗(7)↘(27)	↗(43)↘(9)	↗(35)↘(3)
sig. signs /countries for TFP	↗(27)↘(28)	↗(17)↘(23)	↗(15)↘(19)	↗(9)↘(12)
sig. signs /countries for $Open$	↗(21)↘(20)	↗(16)↘(12)	↗(19)↘(6)	↗(12)↘(5)

Note: Dependent Variables natural logarithm of the Gini coefficient. All variables are in natural logarithms. Values between parentheses below coefficients are p-values from robust (clustered) standard errors. Level of significance: *** for p-value<0.01; **for p-value<0.05;* for p-value<0.1. Wald test is a joint significance test for the regressors. CD-test is a Pesaran (2004) cross-section dependence test on the null of cross-section independence done on the residuals from the regression (p-value presented between parentheses). Stat-test is the Pesaran (2007) unit root test made on the residuals. This test used 3 lags and rejects I(1) means that in all lags the test of unit root rejects. The list of countries that enter in columns (3) and (4) are provided in the Appendix B.

as rich countries. Results are in Table 5.6 and show that the positive effect of human capital on inequality, once it is corrected for uncertainty in data, occurs mainly in rich countries. In these countries a 1% increase in human capital would imply that the corrected Gini coefficient increase from 3.2% to 4%. The fact that the positive effect of human capital in inequality is particularly evident on the group of rich countries is consistent with the skill-biased technical change theory, according to which the increase in human capital stocks should be associated with the adoption of skill-biased technologies, which in turn positively influence the wages of the richest in the economy. This effect may overcome the supply effect and is present mostly in the rich countries (see e.g. Hornstein, Krusel and Violante, 2005: 1306).

Below, we present a set of robustness analysis to evaluate the effect of human capital and TFP on inequality, using the uncertainty-corrected measure of the Gini coefficient.

5.4.4 Robustness: alternative specifications

In the robustness analysis we have implemented slightly modified common correlated effects estimators as suggested in recent literature. We include in regressions one or more further covariates in the form of cross-section averages, which helps to identify the unobserved common factors (in the spirit of Pesaran, Smith and Yamagata, 2013). Moreover, we also follow Chudik and Pesaran (2013) in introducing lags of cross-section averages in order to account for possible

Table 5.6: Inequality, Human Capital, TFP, and Openness (Rich versus Poor countries)

	Rich Sample		Poor Sample	
	(1)	(2)	(3)	(4)
Dependent Variable: Gini Measure	Gini Net post-tax; post-transfer (./sd)	Gini Net post-tax; post-transfer (>30, ./sd)	Gini Net post-tax; post-transfer (./sd)	Gini Net post-tax; post-transfer (>30, ./sd)
$hcap$	4.043***	3.157***	1.169	0.518
	(0.002)	(0.005)	(0.239)	(0.487)
TFP	-0.127	-0.251	-0.041	-0.030
	(0.656)	(0.444)	(0.717)	(0.860)
$Open$	0.032	-0.119	0.001	-0.078
	(0.784)	(0.242)	(0.985)	(0.315)
N Observ.	1657	1431	1643	1162
Avr. N Obs.	36.8	40.9	28.3	35.2
Min-Max	12-52	22-52	7-48	21-48
Number Countries	45	35	58	33
Wald	9.77**	9.98**	1.52	1.52
CD-test (res)	-	-1.40 (0.162)	-	-0.79 (0.430)
Stat-test (res)	-	reject I(1)	-	reject I(1)

Note: Dependent Variables natural logarithm of the Gini coefficient. All variables are in natural logarithms. A constant is included in the regressions but omitted from the Table. Values between parentheses below coefficients are p-values from robust (clustered) standard errors. Level of significance: *** for p-value<0.01; **for p-value<0.05;* for p-value<0.1. Wald test is a joint significance test for the regressors. CD-test is a Pesaran (2004) cross-section dependence test on the null of cross-section independence done on the residuals from the regression (p-value presented between parentheses). Stat-test is the Pesaran (2007) unit root test made on the residuals. This test used 3 lags and rejects I(1) means that in all lags the test of unit root rejects. The list of countries that enter in columns (3) and (4) are provided in the Appendix B. ./(sd) indicates when the Gini coefficient is divided by the source standard-deviation to account for data uncertainty.

feedback effect from inequality to human capital.[16]

To this end, we consider openness as a cross-section average, seeking to identify the unobserved common factors as linked with globalization and global integration (e.g. the entrance of China in global markets affecting all the countries). Column (1) in Table 5.7 presents these results. In column (2) in the same table we present regressions in which we identify the common unobserved factors as, not only globalization and integration (using the variable openness as cross-section average) but also technological spillovers (using the variable TFP as cross-section average). In column (3) we add to the set of possible unobserved common factors, production spillovers, including GDP per capita as a cross-section average. In column (4) we consider only openness as cross-section average and eliminate TFP from the regression. This regression aims to show that the robustness of the positive effect of human capital on inequality is not dependent on the presence of TFP, and thus, not dependent on the way this particular TFP measure is calculated. In columns (5) and (6) we also include lags of the cross-section averages.[17]

In this robustness analysis we consider as dependent variable the Gini coefficient (net definition), using only cross-sections with more than (or equal to) 30 time-series observations. This is done to allow for diagnostic testing. We will also describe the results obtained with the same variable from all the cross-sections (independently of time-series coverage).

In regressions in which production spillovers are not considered as cross-country common factor - columns (1), (2) and (4) - the effect of human capital is highly significant meaning that

[16]This is similar to what Eberhardt and Presbitero (2014) did in an empirical implementation for the relationship between growth and debt.

[17]We closely follow the rule of thumb suggested by Chudik and Pesaran (2013) - $p = T^{1/3}$ - and include 3 to 4 lags of the cross-section averages.

a 1% increase in human capital would imply a rise in the level of inequality that is around 3.8%. From these, columns (1) and (2) present residuals that show no evidence of nonstationarity or cross-country dependence. Regression residuals from column (4) regression present some evidence of cross-country dependence (yet much lower than in the regressors) and no evidence of nonstationarity. In fact, as in Eberhardt and Prebistero (2014), the introduction of additional cross-country averages in regressions helps to obtain cross-country independence of residuals. In the regression that includes production spillovers as a possible common factor - column (3) - the effect of human capital decreases quantitatively but maintains the high level of significance. In this case, a 1% increase in human capital would imply a rise in the level of inequality of around 1.9%. Additionally residuals show no evidence for cross-country dependence or nonstationarity. For regressions robust to potential feedback effect from inequality to human capital - columns (5) and (6) - the effect of human capital is also significantly positive with comparable absolute effects (3.31% and 2.97% respectively) although the statistical significance is decreased from previous regressions. Wald tests point to high significance of the regressors. The majority of countries present significant coefficients for human capital (from 12 to 44, of which not more than 10 are significantly negative). A relatively high number of countries (27) also present significant coefficients for TFP - in column (1) - although in this case there is a balance between significantly positive and significantly negative results. The most significant individual change that occurred in those regressions that abstain from considering openness as a country-specific determinant of inequality, is the entrance of the USA to the set of countries for which a significantly positive effect of human capital occurs, a fact common to all the regressions in Table 5.7.

Regressions that include all the cross-sections (and not only those with high time-series coverage, as those in the Table 5.7) would confirm those results. Regressions corresponding to those in columns (1), (2) and (4) slightly decrease the effect of human capital to a coefficient from 2.8 to 3.17 (with a high significance corresponding to p-values of 0.000). Regression corresponding to that in column (3) decreases the quantitative effect and the level of significance (to a value near 0.8 and a significance level of near 0.25). Regressions corresponding to those in columns (5) and (6) highly increase the statistical significance of the human capital coefficient and also its absolute value, with 4.7% increase in inequality deriving from a 1% change in human capital.

Table 5.7: Inequality, Human Capital, TFP, and Openness (Robustness)

Dependent Variable	Gini Coefficient net income (./sd, >30)					
Vars. only as CS Avr.	Open	Open; TFP	Open; TFP; GDP p.c.	Open; without TFP	TFP	Open; TFP
Lags of CS Avr.	0	0	0	0	3 (TFP); 4 (other)	3 (TFP, Open); 4 (other)
	(1)	(2)	(3)	(4)	(5)	(6)
$hcap$	3.801*** (0.000)	3.854*** (0.000)	1.984*** (0.001)	3.716*** (0.000)	3.312** (0.026)	2.974* (0.075)
TFP	-0.204 (0.248)	-	-	-	-	-
N Observ.	2593	2855	2855	2855	2383	2240
Avr. N Obs. Min-Max	38.1 21-52	38.6 21-52	38.6 21-52	38.6 21-52	32.2 17-48	33.9 24-48
Number Countries	68	74	74	74	74	66
Wald	78.80***	97.11***	75.50***	133.90***	49.13***	36.07***
CD-test (res)	-0.20 (0.839)	0.81 (0.420)	-1.01 (0.314)	1.89* (0.058)	1.12 (0.261)	0.52 (0.602)
Stat-test (res)	Reject I(1)	Reject I(1)	Reject I(1)	Reject I(1)	Reject I(1)	Reject I(1)
sig. signs /countries for $hcap$	↗(37)↘(2)	↗(44)↘(6)	↗(22)↘(9)	↗(42)↘(8)	↗(13)↘(9)	↗(14)↘(6)
sig. signs /countries for TFP	↗(12)↘(15)	-	-	-	-	-

Note: Dependent Variables natural logarithm of the Gini coefficient. All variables are in natural logarithms. A constant is included in the regressions but omitted from the Table. Values between parentheses below coefficients are p-values from robust (clustered) standard errors. Level of significance: *** for p-value<0.01; **for p-value<0.05;* for p-value<0.1. Wald test is a joint significance test for the regressors. CD-test is a Pesaran (2004) cross-section dependence test on the null of cross-section independence done on the residuals from the regression (p-value presented between parentheses). Stat-test is the Pesaran (2007) unit root test made on the residuals. This test used 3 lags and rejects I(1) means that in all lags the test of unit root rejects. The list of countries that enter in columns (3) and (4) are provided in the Appendix B. Vars. only as CS Avr. means variables that enter regressions only as cross-section average but not as country-specific variable. ./sd indicates when the Gini coefficient is divided by the source standard-deviation to account for data uncertainty.

5.4.5 Discussion

In this section we critically discuss our results and also present some information about additional tests that are not presented in the paper but that are available upon request. We present evidence on the effects of human capital, TFP and openness on inequality. To that end, we used a recent measure of inequality with high coverage (Solt, 2009) and also recently developed estimators that allow for country heterogeneity and are robust to country dependence, stationarity and endogeneity toward unobserved common factors (generally described in the survey from Eberhardt and Teal, 2011). We found that there is great heterogeneity concerning the effects of TFP and openness on inequality. Thus, theories that are not based on country heterogeneity to explain the relationship between technology, openness and inequality miss an important part of the story. Institutions and history may be behind those heterogeneous effects. We also found a positive robust effect of human capital on inequality. This does not dismiss that some heterogeneous effects between different countries are also present. However an overwhelming majority of countries present positive effects and the global effect is positive and significant among several different specifications. We also discovered that this influence of higher human capital in higher inequality is totally dependent on correcting the Gini coefficient for its measurement uncertainty (with a measure of uncertainty provided by the source). According to Solt (2009) the provided standard-error for the Gini coefficient aims to correct the remaining uncertainty in the estimations for the inequality measure. This standard-error measures the remaining error due to lack or poorer information available for some country-year pairs. Interestingly, ignoring this correction would yield a negative and significant effect of human capital on inequality, thus implying allegedly that human capital investments would decrease inequality, a result that would be in opposition to the most recent theories of the skill-biased technological change or general purpose technologies. A deep analysis of the data reveals that such a negative sign of the coefficient for the uncorrected Gini index is due to poorer precision in Gini coefficients. For instance, restricting the regression of column (3) in Table 5.5 to values for the Solt (2009) standard-error above the median would yield a significantly negative coefficient of -0.788 (with a p-value of 0.000) and doing the same to the regression of column (4) in the same Table would yield a coefficient of -0.596 (with a p-value of 0.010). Thus, there is a clear need to account for these differences in quality of the source data when assessing the determinants of inequality.

There are two main issues that might compromise our results: (1) the use of a certain measure of human capital and (2) the correction of the Gini measure with the source standard-error to account for different data quality across the world. Would it be possible that this effect is linked with the specific human capital variable used in this paper? In fact, measurement of human capital has always been somewhat controversial in the literature. The measure of human capital that is most used in the literature is that of Barro and Lee (2001), which has been criticized by e.g. Cohen and Soto (2007) due to measurement errors and sources. In fact, Cohen and Soto (2007) argued to have crucially increased the data quality when compared to their predecessors. Barro and Lee (2013), in the version 1.3 of the database, updated the data to incorporate the criticism. The PWT 8.0 human capital variable used in this paper builds on Barro and Lee database, version 1.3. Additionally, the authors of PWT 8.0 filled in the years between the 5 year intervals provided by Barro and Lee, using linear interpolation and corrected the years of schooling to different returns from schooling by level of education following a Mincerian approach. There are, of course, some limitations of this measure, especially the fact that it does not distinguish the returns from schooling by country and by year. An exploration of the returns to schooling variability in a human capital measure would certainly be obtained at the cost of reducing the country coverage and increasing measurement error. Thus, the human capital variable from PWT 8.0 is the human capital data with widest coverage, and thus the only that consistently allow for the use of heterogeneous panel data methods. In order to investigate if the use of returns to obtain the Mincerian-consistent measure of the PWT 8.0, we repeated the regressions in Tables 5.5, 5.6 and 5.7 using two original alternative variables from Barro and Lee (2013), educational attainment above 15 and 25 years (which were linearly interpolated to obtain comparable series to the one used in the benchmark analysis). The results showed very consistently with previous ones, showing a highly statistical significant and positive effect of human capital on inequality for both variables in all specifications. When comparing the obtained results with those of the tables above, we noted that despite the very high statistical significance (almost always with p-values equal to 0.000), coefficients are slightly lower than those presented on the tables, oscillating between 1.3 and 2.6, indicating that a 1% increase in years of schooling imply an increase in inequality from 1.3% to 2.6%. The remaining effect to those reported in the tables above should be attributed to differences in returns throughout the different levels of schooling. In order to investigate whether the interpolation approach would have eliminated the significance of our results, we ran regressions that eliminated the interpolated observations. This greatly decreased the number of observations available for each regression from nearly 3200 observations to nearly 500 observations. Nevertheless, all regressions corresponding to specifications presented earlier in Table 5.7[18] maintain the highly significant positive signed human capital coefficient, with statistical significance of 5% or less.

The human capital variable construction and the very robust results we have obtained give us confidence that the obtained results must be common to any correct measure of human capital given that it has the wide time-series and cross-country coverage as does this one. As a consequence our strong effect of human capital on inequality has non-negligible policy effects. Until now, and given the results in Barro (2000), the common wisdom has been that if some education increases inequality, it should be the higher levels of education. However, by construction, the employed measure of human capital strongly weights lower levels of education (due to higher returns for lower levels of education). Thus, the effect of education on inequality should also be due to lower levels of education. This has policy relevance as politicians should be aware of this effect in promoting education, even at the lower levels. Notwithstanding, this effect is absent from the poorer countries, which indicates no influence of education in increasing inequality on those countries. Thus, generally, in poorer countries, policy may enhance education with no caution about rising inequality.

The second issue is related to the correction of the Gini coefficient. We did that by simply dividing the Gini coefficient by the standard-error, as explained above. This standard-error oscillates in the sample from 0.0016 to 15.43, which gives an idea of the difference in quality remaining in data and suggests the need to account for these quality heterogeneity. In fact, 25% of the observations present a standard-error below 0.5. Dividing the Gini coefficient by this standard-error would greatly magnify Gini coefficients in the case of high precision (i.e. when standard-deviations approach zero). A correction that would not present that property would be the division of the Gini coefficient by (1+standard-error).[19] With this, a high precision Gini

[18]Considering specifications in columns (1) to (4), as the time-series requirements of specifications in columns (5) and (6) are not met when considering only five-year periods.

[19]The alternative proposed uncertainty-corrected measure is thus $\frac{GINI}{1+sd(GINI)}$, where *GINI* is the Gini index provided by SWIID and *sd(GINI)* is the standard-deviation of the Gini index, also provided by the SWIID and that corrects for uncertainty or measurement error within the sources. Results are provided in Appendix C.

coefficient - with a standard-error close to 0 - would not be increased although a low precision coefficient would be decreased. The high significance of human capital positive coefficients hardly changes with this modification in the corrected Gini index in all the different specifications we present in the paper (corresponding to specifications in Tables 5.5 - columns (3) and (4)- Tables 5.6 and 5.7). The only expected difference in results is quantitative (see Tables D.3.1, D.3.2 and D.3.3). With this alternative variable, a 1% increase in human capital would increase inequality by between 0.62% to 1.52% (compared to 1.98% to 3.85% with the baseline measure). The causal relationship between human capital and inequality in regressions corresponding to specifications in Table 5.7, but in which all the cross-sections (and not restricted to the ones with larger time-series) are included, is also robust to the mentioned change in the definition of the corrected Gini coefficient. The original variables from Barro and Lee (2013), for educational attainment above 15 and 25 years, present also a robust influence in inequality if the measure of inequality changes according to the described above (i.e. dividing the Gini coefficient by (1+standard-error)).

5.5 Conclusion

There is scarce quantitative literature on the determinants of inequality. We contribute to that literature by evaluating potential determinants of inequality in a large panel data of countries. Earlier attempts have faced problems with the coverage and quality of the income inequality data. We use a recent standardized measure of the Gini coefficient, due to Solt (2009) to evaluate human capital, TFP and openness as possible determinants of inequality. We conclude that this measure also needs to be corrected for differences in original data precision. Failure to do so would determine crucially different and misleading results concerning the influence of human capital on inequality. Fortunately, Solt (2009) also provides the means to implement such correction.

We found that inequality data, as well as other macroeconomic variables, are subject to cross-country dependence and nonstationarity and so, newly developed econometric methods designed to analyze moderate T, moderate N panels should be employed (Eberhardt and Teal, 2010a). We proceeded along this line and implemented cointegration tests to evaluate the (Granger) causality between human capital and inequality. Results indicate a strong channel from human capital to inequality.

Regressions based on heterogeneous panels methods indicate that there is great heterogeneity concerning the effects of TFP and openness on inequality. Additionally, we found a positive statistically significant effect of human capital on inequality once the Gini coefficient is corrected for differences in it«s precision. This result is robust to several specification changes and measurement changes both in the inequality variable and in the human capital variable. Notably, the positive effect of human capital on inequality remains highly significant in methods robust to reverse causality. This does not dismiss that some heterogeneous effects between different countries are also present. Contrary to what may have been the current wisdom until now, it is not only tertiary education that tends to cause higher inequality, but the effect is highlighted with a measure that strongly weights lower levels of education, suggesting further research on the effect of primary education on inequality.

These results suggest that theories that are not based on country heterogeneity to explain the relationship between technology, openness, and inequality may be unrealistic. In fact, institutions and history may be behind the heterogeneous effects of human capital, technology, and openness on inequality detected. Additionally, contrary to earlier evidence, the results in this paper suggest that human capital may be seen as the most important worldwide determinant of inequality.

Chapter 6

Income Inequality and Technological Adoption

6.1 Introduction

A strong and active theoretical literature seeks to explain the rise of income inequality in the second-half of the twentieth century alongside the rise in the supply of human capital. Skill-biased technical change and capital-skill complementarity are crucial to explain these phenomena. Generally, according to theory, skill-premia increase due to two effects. First, the skill premium would reflect the productivity difference between sectors. Second, with full capital mobility, factor price equalization requires capital to flow to the sector operating with the new technology. Workers in the new technologies sectors are thus endowed with more capital, which raises their relative wages (Acemoglu, 2002a, 2002b, 2003).

An alternative development has argued that the diffusion of IT - General Purpose Technologies - may have increased the demand for adaptable skilled workers and made vintages of capital more adaptable. This in turn increases the premium of workers that show a lower learning cost and that can adapt quickly from one sector to another. These ideas have been formalized by Galor and Tsidon (1997), Greewood and Yorukoglu (1997), Caselli (1999), Galor and Moav (2000), and Aghion *et al.* (2002).

Whatever the explanation may be for the rise in income inequality and its relationship with technology, there is very little quantitative literature on the issue, as observed by Hornstein *et al.* (2005:1361). In fact, empirical attempts to evaluate the relationship are mostly country-specific as are Ding *et al.* (2011) and Rattsø and Stokke (2013), for example. Barro (2000) and Jaumotte *et al.*(2013) examined this relationship in large samples of countries. Barro (2000) presents fixed-effects estimations of equations of the Gini index on covariates such as GDP and GDP squared, schooling, democracy index, openness, rule of law index, and several dummies. In his fixed-effects estimations, dummies for income or spending, secondary and higher education are negatively related to inequality and openness is positively related to inequality. Primary schooling and the dummy for individual or household data are insignificantly related to the Gini coefficient. There is a strong inverted-U relationship with GDP: the so-called Kuznets curve. Recently, Jaumotte *et al.* (2013) have re-assessed the determinants of inequality. They focused on the effect of globalization on inequality but do not go into the relationship between inequality and GDP. They concluded that trade globalization decreases inequality while financial globalization increases inequality. Moreover, information and communication technologies and credit deepening increases inequality while the share of industry in the economy decreases it. Education variables and initial GDP (when included) are insignificantly related to inequality. The evidence relating different types of technologies and inequality, as far as we know, does not exist. We contribute to fill this gap.

This paper's contribution is twofold: first, it uses a large dataset on technology adoption (from Comin and Hobijn, 2009) to evaluate their effect on income inequality; second, it evaluates the effects of different technologies on inequality taking into account country heterogeneity, cross-country dependence and endogeneity to common factors. We are thus able to identify which technologies are most equality-friendly or inequality-friendly and with this we highlight some new evidence. In particular, we are able to evaluate for the first time the effect of the adoption of some individual technologies (such as tractors, TV, aviation, railways, etc.) whose effect has been neglected in the study of inequality. To this end, we have also obtained measures of aggregate technology adoption by type of technology - modern ICT, older ICT, production, and transport technologies. Our main conclusions point to a positive effect of older ICT technologies (includes radios, mainline telephone lines, televisions in use, and telegrams) and transport technologies (includes aviation, railways lines, steamships, passenger cars, and commercial vehicles), and to a lesser extent of modern ICT technologies (includes computers, ATMs, internet users, and cell phones). Our results also indicate that the effects of technology adoption may be quite different from country to country and from groups of rich and poor countries.

The remainder of the paper is organized as follows. Next, in Section 6.2 we describe our data set and respective sources. In Section 6.3 we describe our estimation strategy. In Section 6.4 we describe our results using estimators robust to country heterogeneity, cross-dependence, and endogeneity. Section 6.5 concludes.

6.2 Data and Sources

There are currently three different projects that collect and make publicly available inequality data for many countries and periods around the world: the Luxembourg Income Study (LIS), the data set assembled by Deininger and Squire (1996) for the World Bank (WIID), recently updated and upgraded by the WIDER (World Institute for Development Economic Analysis) project, and the most recent standardized World Income Inequality data set (SWIID), by Solt (2009). The LIS, which was used by Jaumotte *et al.* (2013), has generated the most-comparable income inequality statistics currently available but covers relatively few countries and years. The Deininger and Squire data set and its successors, used by Barro (2000), on the other hand, can be used to provide many more observations, but only at a substantial loss of comparability. Solt (2009) implemented a sequence of steps to standardize income inequality data and provide data with wider coverage than the WIID, but at the highest quality as in LIS. However, in the process of standardization, not all countries had the sufficient data in the original sources. To handle this, Solt (2009) also calculated a standard-error of each Gini coefficient to account for the remaining uncertainty in data. We use data for inequality from the Standardized World Income Inequality database (SWIID), version 4.0, from Solt (2009), for the Gini coefficient and for the respective standard-error.[1] This includes data on the Gini coefficient, using post-taxes, post-transfers income (the net definition) and on the Gini coefficient, using pre-taxes, pre-transfers income (the market definition), and the respective standard-errors by country and year. We selected the net definition of the Gini coefficient as it accounts for the distortionary effects that fiscal systems have on income distribution of countries. Our measure of uncertainty-corrected Gini index is the following:

$$Gini_{i,t} = \frac{GINI_{i,t}}{1 + sd(GINI)_{i,t}} \quad (6.1)$$

[1]Available at $http://thedata.harvard.edu/dvn/dv/fsolt/faces/study/StudyPage.xhtml?studyId = 36908$.

where $Gini_{i,t}$ is the net definition of the Gini index given by the 4.0 version of the SWIID and $sd(GINI)_{i,t}$ is the standard-deviation of the net definition of the Gini index given by the 4.0 version of the SWIID, which measures the uncertainty or measurement error of the Gini index.

For technologies adoption we use the CHAT database from Comin and Hobijn (2009) and concentrate on the 20 largest technologies as used in Comin *et al.* (2013). First, we will present results on the effect of individual technologies on inequality. For each measure, and inspired in the theory that relates skill-technological complementarities with inequality (Acemoglu, 1998), we consider a measure of skill-technological complementarity for each pair country (*i*), year (*t*), such as:

$$Techh_{j,i,t} = tech_{j,i,t} \times hc_{i,t} \quad (6.2)$$

where $Techh_{j,i,t}$ is our measure of technology adoption (considering skill-technological complementarity, $tech_{j,i,t}$ is the natural logarithm of one of the j technology adoption measures coming from Comin and Hobijn (2009), and $hc_{i,t}$ is the natural logarithm of the human capital measure coming from the Penn World Tables 8.0. We also use as an additional control variable, which may influence inequality, the log of the Openness ratio from the Penn World Tables 8.0. Education variables and openness variables are also in the earlier articles that studied the determinants of inequality in a large cross-section of countries (Barro, 2000 and Jaumotte *et al.* (2013) and so, we will also include them in our regressions.

Below, in order to summarize information about technologies adoption, we create additional variables such as ICT(modern), Transportation, Production and ICT (older). Each of these are sums of standardized values of technologies $Techh_{j,i,t}$ that lie in each category. In the modern ICT we included computers, ATMs, internet users, and cell phones. In the Transportation we included civil aviation passenger-kilometers traveled, civil aviation ton-kilometers traveled, public railway lines, passenger journeys by railway in passenger-kilometer, freight carried on railways (excluding livestock and passenger baggage), steamships, passenger cars and commercial vehicles. In the Production technologies we include wheel and crawler tractors, (excluding garden tractors), gross output of electric energy (inclusive of electricity consumed in power stations) in Kw-Hr, crude steel production (in metric tons) in blast oxygen furnaces and crude steel production (in metric tons) in electric arc furnaces. For the older ICT we included number of radios, mainline telephone lines, number of televisions in use and telegram. For each of the constructed technologies types, and in order to maximize the time-series coverage, we considered that each sum includes values when at least one of its components has values in each country-year pair. Any missing value is also taken as evidence of no technology adoption. We will also discuss results with alternative assumptions that, of course, come at the cost of lower coverage.[2]

We end up with an unbalanced panel database of a maximum of 111 countries with a minimum of 1 year *per* country and a maximum of 42 years *per* country. The initial year covered is 1960 and the last 2003. These values depend on the technology considered. Among the technologies with excellent coverage in the database, we count electrical production, tractors, rail line, telephone, TV, and vehicles. On the contrary ATMs, internet, ships and steel are among the less covered. Coverage oscillates between 368 observations (ATMs) to 5991 observations (electrical

[2]Had we restricted the technology-types measures to sums in which all the parcels had non-missing values, the resulting number of observations would be insufficient to perform regressions with the four types as regressors.

production). Table 6.1 shows descriptive statistics for all the variables included in the analysis. Details for definitions and sources are given in the Appendix.

Table 6.1: Descriptive statistics

Variable	Obs	Avr	S.d.	Min	Max
ag_tractor (tt)	5190	163183.1	535339.5	2	5470000
atm (atm)	368	18318.39	43956.15	22.608	370782.8
aviationpkm (a1)	3535	7529.996	42396.95	0	772000
aviationtkm (a2)	3157	224.7141	904.3084	0	14788
cellphone (cp)	3963	1046050	7260919	0	2.06E+08
computer (ct)	1350	2943427	1.24E+07	4.097402	1.90E+08
elecprod (ep)	5991	5.27E+10	2.06E+11	100000	3.20E+12
internetuser (it)	1446	1753685	9086197	0	1.59E+08
radio (ra)	5614	10305.62	43871.5	0	585000
railline (r1)	4584	11939.46	35181.85	0	361049
railpkm (r2)	3305	16487.33	51290.82	0	414000
railtkm (r3)	3667	58422.92	306202.9	0	3900000
shipton_steammotor (sh)	1752	3778.575	9293.092	7	81528
steel_bof (s1)	1412	9040.385	15971.17	4	100000
steel_eaf (s2)	2212	2714.221	5698.591	1	47850
telegram (tg)	2466	12.64869	24.44503	0	312.24
telephone (tl)	5255	3028552	1.40E+07	300	2.14E+08
tv (tv)	4728	5836445	2.50E+07	10	4.12E+08
vehicle_car (cr)	5095	2591432	1.32E+07	100	2.22E+08
vehicle_com (tr)	4710	725.3957	4750.137	0.1	88000
modern_techs_st_aug	10899	0	1.00174	-6.422771	12.23062
comm_techs_st_aug	10899	0	1.900109	-5.824181	12.35615
transport_techs_st_aug	10899	0	3.373964	-9.800905	32.98412
prod_techs_st_aug	10899	0	1.657976	-6.283744	11.1706
hc	6694	2.093233	0.6318047	1.018154	3.618748
Open	7760	0.4888119	0.6575676	2.93E-06	24.68241
lgini_st1	4597	2.747416	0.4281077	1.194536	3.80669

6.3 Estimation and Methods

Our baseline specification is as follows:

$$gini_{it} = \beta_1 Techh_{jit} + \beta_2 hc_{it} + \beta_3 Open_{it} + \lambda_i' \mathbf{f}_t + u_{it} \quad (6.3)$$

where $gini$ is the natural logarithm of the uncertainty-corrected Gini coefficient given by (6.1), $Techh_{jit}$ is the measure of technology adoption given by (6.2), hc_{it} is the measure of human capital, and $Open_{it}$ is the measure of openness, all described above. Thus the coefficient on our measure of technology β_1 measures the effect that a skill-complementary technology has on inequality or, in other words, the effect of technology on inequality that depends on the existing level of human capital. A positive coefficient means that a higher level of adoption of a given technology or a given type of technology causes a higher level of inequality, an influence that is dependent on human capital. Thus, higher levels of human capital enhance the effect of a given technology on inequality. If the coefficient is negative, the effect of the skill-complementary technology or type of technology tends to decrease inequality, which indicates that a higher level of adoption decreases inequality and this negative effect is enhanced by the existing level of human capital. This effect may be conditional on a direct effect of human capital, captured by β_2 . Finally, we may consider that technology adoption is being determined by the same phenomena as inequality, by common factors such as globalization or the entry of China into the world market and technology would thus become an endogenous variable. These common factors are accounted for in $\mathbf{f_t}$.[3] λ_i' is the vector of factor loadings associated with the common factors. As can be observed from equation (6.3) each coefficient is country-specific, thereby allowing for complete heterogeneity in the estimation. Additionally, as each regressor can also depend on the common factor, the method is in fact robust to endogeneity of the observable factors toward the common factors determining inequality. The estimation is performed using the Pesaran (2006) common factor estimator in the baseline analysis. As Pesaran and Tosetti (2011) explain, this method is robust to non-stationarity in both observable and non-observable variables and works well in the presence of weak and/or strong cross-sectionally correlated errors.

6.4 Results

In this section we begin by presenting and analyzing results for the influence of each technology on income inequality and then the results for the influence of the four technology-type measures.

6.4.1 Influence of 20 Different Technologies on Income Inequality

In this section we present regressions with specification (6.3) in which $Techh_{jit}$ assumes each of the 20 technologies considered by Comin *et al.* (2013).

In order to allow for a comparison with results without skill-complementary technologies, we first describe the results of a regression without them. In a regression in which only human capital (*hc*) and openness were considered (i.e. restricting $\beta_1 = 0$), human capital would be highly significant (p-value of 0.000) with a coefficient of 1.2, meaning that a change in human capital of 1% would raise the inequality index by 1.2%. Openness however would present a less significant effect of 0.032, with a significance level of only 19.5%. This regression would include 123 countries with a minimum of 4 and a maximum of 52 observations. The Wald test indicates the global significance of regressors.

Tables 6.2 and 6.3 present the results. Results indicate positive and significant effects of aviation, cell phones, electric production, internet, telephone and TV on income inequality. This broadly confirms the theoretical results according to which ICT and general purpose technologies (such as aviation and electricity) adoption tend to increase inequality. Quantitatively, significant elasticities are between 0.03 (cellphones) to 0.22 (telephones), meaning, e.g., that a 1% increase in the use of telephones will imply an increase in inequality of 0.22%, for a given level of human capital. Interestingly the introduction of skill-complementary technologies in the regressions implies that a direct effect of human capital almost disappears, with only four exceptions, columns (3), (5), (6), and (10) in Table 6.3.

[3]For complete arguments toward reconsideration of traditional econometric methods to study moderate-T dimensional panel data of countries, see Eberhardt and Teal (2011).

Table 6.2: Inequality and Technologies: Part I

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$Techh$ def.	tt	atm	a1	a2	cp	ct	ep	it	ra	r1
$Techh$	0.02	0.10	0.05	0.07*	0.03**	0.02	0.14**	0.07***	0.11	-0.14
	(0.89)	(0.30)	(0.38)	(0.07)	(0.04)	(0.65)	(0.03)	(0.00)	(0.44)	(0.65)
hc	-0.85	-2.33	0.34	0.68	0.95	1.22	-2.36	-1.06	-0.86	1.96
	(0.49)	(0.38)	(0.66)	(0.20)	(0.21)	(0.38)	(0.17)	(0.52)	(0.54)	(0.32)
$Open$	-0.01	-0.03	0.07	0.06	0.05	0.03	0.03	0.07	-0.01	-0.04
	(0.66)	(0.89)	(0.18)	(0.24)	(0.41)	(0.55)	(0.29)	(0.47)	(0.73)	(0.47)
Wald	0.69	1.84	2.79	6.37*	0.07*	1.32	7.84**	11.1**	1.10	1.73
Avr. Obs.	22.6	11.2	20.8	19.5	12.4	11.5	22.4	10.0	21.1	23.2
N. Countries	109	33	68	64	106	100	104	107	109	74
Total Obs.	2465	368	1415	1250	1314	1153	2326	1075	2304	1718

Note: Dependent Variables natural logarithm of the Gini coefficient. Definitions: tt - tractors; atm - ATM machines; a1 - aviation passengers; a2 - aviation cargo; cp - cellphones; ct-computers; ep- electricity production; ra- radios; r1 - length of rail lines. A constant is included in regressions but omitted from the table. Values in parentheses are p-values. P-values of coefficients are based on robust standard errors. Level of significance: *** for p-value<0.01; **for p-value<0.05;* for p-value<0.1. Lists of countries included in individual regressions are available upon request.

Table 6.3: Inequality and Technologies: Part II

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$Techh$ def.	r2	r3	sh	s1	s2	tg	tl	tv	cr	tr
$Techh$	0.08	-0.04	0.06	0.03	0.00	-0.09	0.22**	0.18**	0.11	0.06
	(0.39)	(0.52)	(0.67)	(0.30)	(0.90)	(0.15)	(0.05)	(0.02)	(0.20)	(0.35)
hc	1.04	0.87	1.67*	0.85	1.13*	1.80***	-0.60	-1.33	0.52	1.49**
	(0.19)	(0.17)	(0.07)	(0.15)	(0.08)	(0.0)	(0.76)	(0.30)	(0.64)	(0.02)
$Open$	0.09	-0.00	-0.07	-0.02	0.12***	0.02	0.06	-0.01	0.03	0.03
	(0.10)	(0.95)	(0.25)	(0.77)	(0.01)	(0.54)	(0.11)	(0.78)	(0.48)	(0.52)
Wald	5.12	2.26	4.97	3.20	9.86**	12.1***	6.72*	7.23*	2.35	6.83*
Avr. Obs.	20.8	21.1	21.6	23.9	22.9	20.8	18.5	21.6	20.6	22.6
N. Countries	59	64	39	47	71	49	104	111	99	77
Total Obs.	1229	1353	843	1124	1624	1021	1929	2402	2038	1742

Note: Dependent Variables natural logarithm of the Gini coefficient. Definitions: r2 - railway passengers; r3 - railway cargo; sh - ships; s1 - steel in blast oxygen furnaces; s2 - steel in electric arc furnaces; tg - telegrams; tl- telephones; tv- televisions; cr- passenger vehicles; tr - commercial vehicles. A constant is included in regressions but omitted from the table. Values in parentheses are p-values. P-values of coefficients are based on robust standard errors. Level of significance: *** for p-value<0.01; **for p-value<0.05;* for p-value<0.1. Lists of countries included in individual regressions are available upon request.

6.4.2 Influence of Technology Types

In order to summarize results we built a taxonomy of four technology types, as described above. We now present the results for the influence of those technology types on inequality. This also allows us to analyze the conditional effect of each technology type on inequality, which enables answering the question if inequality rises due to e.g. ICT for the same adoption of other technology types. Thus, $Techh_{jit}$ now assumes one of the four types: modern ICT, older ICT, production, or transportation.

We continue to employ Pesaran (2006) common correlated effects estimator but we also implement slightly modified common correlated effects estimators suggested in recent literature. We include in the regressions one or more additional covariates in the form of cross-section averages, which helps to identify the unobserved common factors (in the spirit of Pesaran *et al.*, 2013 and following what Eberhardt and Presbitero, 2014 did in an empirical implementation). To this end, we consider openness as a cross-section average, seeking to identify the unobserved common factors as linked with globalization and global integration (e.g. the entrance of China in global markets affecting all the countries). In some of the regressions, together with openness we also considered averaged TFP as an additional control. This is to identify the unobserved common factors also with productivity spillovers around the world. Given that we now have available several more time-series observations *per* country, we present specifications with country trends, as well as information on their significance across countries. Table 6.4 presents these results.

Results indicate a highly significant effect of transportation technologies adoption on the increase of inequality, conditional on the adoption of other technology types. Thus, it seems that countries that adopted more transportation technologies than other technologies have also faced, due to that, an increase in inequality. This is a somewhat unexpected result, as theory has focused more on information and communication technologies as a source of inequality. Countries that adopted more transportation technology may be highly integrated in world trade and thus be highly competitive. This can influence the wages of the most adaptable workers and thus increase inequality. In fact, transportation technologies are general purpose technologies in the sense that they are applied to the economy as a whole, with important effects on sectoral and countries integration. Results also reveal positive effects of older ICT technologies (specifications (2) and (4)) and of production technologies (specification (1)). Curiously, modern ICT has a non-significant effect on inequality, conditional on other technology-types adoption. Quantitatively, the effects mean that a 1 standard-deviation (s.d.) increase in a skill-complementary transportation technology (s.d.=3.37) would increase inequality by 3.37% to 6.74%. If the 1 standard-deviation rise occurs in the older ICT technologies (when significant), for a s.d. equal to 1.90, the implied rise in inequality will amount to a value between 1.90% and 3.80%. Finally, If the 1 standard-deviation rise occurs in the production technologies (when significant), for a s.d. equal to 1.66, the implied rise in inequality will amount to 3.32%.

It is interesting to evaluate if these results differ from rich countries to poor countries, even before we analyze effects by individual country. Several features that could influence the relationship between skill-complementary technologies and inequality are quite different between rich and poor countries. The level of education, the composition between general and vocational education, and the proximity to the technology leader are only some of them. We next present results in which we divided the sample by the average GDP *per capita*, after averaging GDP *per capita* inside each country. Results shown in Table 6.5 highlight that the robust effects described above are all due to very strong positive effects of the technology types adoption

Table 6.4: Inequality and Technology-types

	(1)	(2)	(3)	(4)	(5)	(6)
Modern ICT	0.00	0.00	0.00	-0.00	-0.00	-0.00
	(0.895)	(0.494)	(0.931)	(0.740)	(0.935)	(0.905)
Older ICT	0.01	0.01**	0.01	0.02*	0.01	0.01
	(0.223)	(0.032)	(0.223)	(0.089)	(0.417)	(0.395)
Production	0.02***	0.01	0.01	0.01	0.00	-0.01
	(0.008)	(0.172)	(0.369)	(0.365)	(0.812)	(0.568)
Transportation	0.01*	0.01**	0.02***	0.02***	0.02***	0.02***
	(0.052)	(0.016)	(0.001)	(0.003)	(0.004)	(0.007)
Trend	-	-0.01***	-	0.01**	-	0.00
		(0.009)		(0.047)		(0.988)
Additional CS Avg	No	No	Open	Open	Open; TFP	Open; TFP
% sig. trends	-	32.3%	-	25.9%	-	27.5%
Wald	9.12*	12.81**	92.04***	80.86***	76.41**	56.93***
Avr. Obs.	29.2	29.4	32.2	32.7	32.7	33.3
N. Countries	156	155	115	112	112	109
Total Obs.	4558	4552	3702	3666	3666	3627

Note: Dependent Variables natural logarithm of the Gini coefficient. A constant is included in regressions but omitted from the Table. Values between parentheses are p-values. P-values on coefficients are based on robust standard errors. Level of significance: *** for p-value<0.01; **for p-value<0.05;* for p-value<0.1. Additional CS Avg means the additional variables added as controls as cross-section averages. % sig. trends is the percentage of country-trends that are statistically significant at the 5% level. Lists of countries included in individual regressions are available upon request.

that occur in rich countries. In fact, in rich countries, both transportation technologies and old ICT adoption are associated with high inequality and also modern ICT causes inequality when a trend is considered (column (2)), confirming the theory result and existing evidence relating ICT to inequality (e.g. in Jaumotte *et al.* (2013)). In the sample of the poorest countries, it is not possible to identify any significant effect of skill-complementary technology on inequality.

6.4.3 Robustness

We have also tested our results against differences in the implemented estimator and in a restricted sample. The alternative estimator was developed by Eberhardt and Teal (2010b), seeking to identify the common unobserved effects with a single common factor designed to estimate a residual such as TFP. The restricted sample is one with higher populated time-series in which we restricted the sample to the countries that had more than 15 time-series observations for the dependent variable.

Generally, the robustness analysis presented in Table 6.6 confirms our previous results. Transportation technology increases inequality throughout all the considered specifications with similar quantitative effects as those obtained previously. Additionally older ICT also contributes to the rise in inequality in specifications in which a (statistically significant) trend is introduced in the regression. The specifications based on the Eberhardt and Teal (2010b) estimator tend to increase the positive effect of modern ICT and in the specification with a (statistically significant) trend - column (4) - it also becomes highly significant. We have also divided the sample between rich and poor countries as we did before and use the Eberhardt and Teal (2010b) estimator.[4] Besides the significant effects obtained in the rich countries sample presented in Table 6.5, we have also obtained for the poor countries positive and highly significant (2.2% and 6.9% levels of significance respectively) coefficients for new and old ICT.

[4]These results are not presented but are available upon request.

Table 6.5: Inequality and Technology-types: Rich and Poor Countries

	(1)	(2)	(3)	(4)
	Rich		Poor	
Modern ICT	0.01	0.02**	-0.00	0.00
	(0.310)	(0.042)	(0.711)	(1.000)
Older ICT	0.01**	0.01**	-0.00	0.00
	(0.024)	(0.023)	(0.787)	(0.676)
Production	0.00	0.00	-0.00	0.01
	(0.969)	(0.980)	(0.943)	(0.307)
Transportation	0.01**	0.01***	-0.00	-0.00
	(0.019)	(0.009)	(0.601)	(0.911)
Trend	-	0.00	-	-0.01***
		(0.166)		(0.004)
% sig. trends	-	33.8%	-	25.6%
Wald	11.59**	15.51**	0.49	1.23
Avr. Obs.	33.4	33.8	26.1	26.1
N. Countries	66	65	90	90
Total Obs.	2205	2199	2353	2353

Note: Dependent Variables natural logarithm of the Gini coefficient. Values in parentheses are p-values. A constant is included in regressions but omitted from the table. P-values of coefficients are based on robust standard errors. Level of significance: *** for p-value<0.01; **for p-value<0.05;* for p-value<0.1. Additional CS Avg means the additional variables added as controls as cross-section averages. % sig. trends is the percentage of country-trends that are statistically significant at the 5% level.

We performed an additional test on all results (which we do not show but are available upon request). Until now our measure of skill-complementary technology is a measure affected by scale, i.e., technological adoption is taken as total technological adoption, thus being influenced by the size of the country. This does not raise any particular problem *de per si*, since the inequality index is independent of the size of the country as well as human capital and openness. The conclusion is that some of those skill-complementary technology adoption affected by scale tend to influence inequality positively. Does an alternative *per capita* skill-complementary technology adoption measure which would be scale-independent have the same effect on inequality? We re-ran all the regressions in the paper with these alternative measures (consisting of dividing the measure in (6.2) by the population in each year and country). Conclusions are as follows:

- Cellphone, internet, telephone, and TV adoption cause more inequality in specifications similar to those in Tables 6.2 and 6.3;
- Telegraph adoption contributes to decrease inequality in specifications similar to those in Tables 6.2 and 6.3;
- Transportation, production, and older ICT are still significant as determinants of (more) inequality in several regressions specified as in Table 6.4;
- Transportation technologies and modern ICT are still responsible for higher inequality in rich countries, as specified in Table 6.5;
- Production technologies and older ICT are significant determinants in the Pesaran (2006) specification, in specifications similar to those in Table 6.6;
- Older ICT is the only significant determinant of higher inequality in the Eberhardt and Teal (2010b) specification, in a specification similar to that in Table 6.6.

Thus, surprisingly, despite a complete re-definition of the relevant measures for skill-complementary technologies adoption, removing the scale dimension of variables, results are quite

Table 6.6: Inequality and Technology-types: Robustness Analysis

	(1)	(2)	(3)	(4)	(5)	(6)
Estimator	Eberhardt and Teal (2010b)	Eberhardt and Teal (2010b)	Eberhardt and Teal (2010b)	Eberhardt and Teal (2010b)	Pesaran (2006)	Pesaran (2006)
Sample	Total		Highly Populated Time-Series (≥ 15)			
Modern ICT	0.00	0.00	0.00	0.01**	-0.00	0.00
	(0.571)	(0.126)	(0.469)	(0.050)	(0.801)	(0.881)
Older ICT	0.00	0.01***	0.01	0.01***	0.01	0.02**
	(0.450)	(0.005)	(0.296)	(0.004)	(0.160)	(0.026)
Production	-0.00	-0.00	-0.01	0.00	0.02*	0.01
	(0.262)	(0.800)	(0.135)	(0.968)	(0.092)	(0.273)
Transportation	0.01*	0.01**	0.01**	0.01**	0.01*	0.01**
	(0.082)	(0.039)	(0.039)	(0.023)	(0.072)	(0.037)
Trend	-	0.00**	-	0.00***	-	-0.00*
		(0.021)		(0.009)		(0.099)
$\%$ sig. trends	-	42.4%	-	44.3%	-	35.1%
Wald	5.17	14.48***	8.13*	17.39***	8.12*	10.52**
Avr. Obs.	29.4	30.0	32.8	32.8	32.8	32.8
N. Countries	155	151	131	131	131	131
Total Obs.	4552	4524	4301	4301	4301	4301

Note: Dependent Variables natural logarithm of the Gini coefficient. A constant is included in regressions but omitted from the table. Values in parentheses are p-values. P-values of coefficients are based on robust standard errors. Level of significance: *** for p-value<0.01; **for p-value<0.05;* for p-value<0.1. Additional CS Avg means the additional variables added as controls as cross-section averages. $\%$ sig. trends is the percentage of country-trends that are statistically significant at the 5% level.

consistent to those obtained when using the scale affected measure of skill-complementary technology adoption. When technology-types measures are considered, production technologies become relatively more important in explaining higher levels of inequality than before, maintaining the relative importance of transportation, old, and modern ICT.

Finally, we wish to discuss a possible alternative to construct the measures of technology types. Alternatively to what has been done, we could have restricted the technology types measures on observations for which all the components presented non-missing values (see Table E.2.1). Due to the fact that this sample is highly unbalanced, doing so would imply very few observations for the technology types variables which would imply that a regression with the four variables as regressors would simply not be possible. However, it would be possible to evaluate the (non-conditional) effect of each technology type to evaluate if, for example, it is crucially different from the conditional one. Interestingly, when testing individually the technology-types variables (restricted to non-missing observations in all the components), despite a huge drop in the number of observations (from nearly 4500 as in Table 6.6 to around 500), all the technology-types have highly significant and positive coefficients (at 1% level), with coefficients oscillating from nearly 0.06 for all technology-types except older ICT type, and 0.14 for older ICT type, highlighting also positive and significant effects for those restricted measures for technology-types.

6.4.4 Results by Country

This section reports the results by individual country. For this we consider the restricted sample with high populated time-series considered also in Table 6.6. This is done in order to maximize time-series availability within countries. Tables 6.7, 6.8, and 6.9 show the statistically significant results by country.

These results highlight the great heterogeneity of the effects of technology adoption on in-

Table 6.7: Countries statistically significant - Part 1

Countries	modern ICT	old ICT	transportation	production
Argentina				0.10 (0.000)
Armenia		-0.27(0.000)		
Australia			0.03(0.007)	
Bolivia	0.29(0.005)	0.55(0.018)	-1.51(0.000)	4.866(0.003)
Botswana	-0.20(0.090)			
Brazil			0.17(0.094)	
Bulgaria		0.19(0.001)		-0.37(0.000)
Burkina Faso				
Burundi	-0.03(0.001)		0.09(0.000)	0.15(0.000)
Cambodia		-0.94(0.057)		
Cameroon				1.54(0.003)
Canada		-0.06(0.034)		0.08(0.000)
Chile			0.22(0.000)	-0.20(0.003)
China			0.07(0.008)	0.16(0.020)
Costa Rica		-0.57(0.030)		
Czech Republic	0.11(0.006)	0.06(0.099)	0.03(0.003)	-0.10(0.068)
Dominican Republic			0.46(0.072)	
Ecuador	-0.14(0.087)			
Egypt			0.14(0.008)	-0.07(0.045)
Estonia	0.33(0.004)			
France		0.11(0.015)	-0.03(0.030)	
Ghana			0.32(0.019)	
Guatemala			0.11(0.001)	
Haiti			-2.04(0.017)	
Hong Kong			0.14(0.039)	-0.26(0.096)

come inequality by country. Despite some globally non-significant signs, there might be a wide range of countries in which technological adoption tends to decrease or increase inequality. Additionally, despite the overall positive signs of the skill-complementary technological adoption coefficients, there might be some countries in which there is evidence that technological adoption decreases inequality.

Of the 131 countries entering in regressions (see e.g. columns (3) to (6) in Table 6.6), there are 75 with significant coefficients in at least one technology-type. As expected, most of the countries present positive signs, i.e., technology adoption causes higher inequality. There are 11 countries in which modern ICT adoption tends to raise inequality and 8 in which this technology type tends to decrease inequality. Among the first group, we find countries such as Netherlands, Iceland, United Kingdom, Thailand and Indonesia. Among the second, we identify countries like Switzerland, Japan, Burundi and Equador. There are many more countries with significant coefficients associated with other technology types than to modern ICT (29 to old ICT and transportation type and 34 to production type). There are 22 countries in which old ICT adoption tends to raise inequality and 7 in which this technology type tends to decrease inequality. Among the first group, we find countries such as France, Netherlands, Singapore, Sweden, Iran, Panama, Mali, and Malawi, to give some examples. Among the second, we identify countries such as Canada, Latvia, Cambodia and Costa Rica. There are 24 countries in which transportation technology-type adoption tends to raise inequality and 6 in which this technology type tends to decrease inequality. Among the first group, we find countries such as Australia, Iceland, Ireland, Burundi, China, Tanzania, and Thailand. The second group comprises Bolivia, France, Haiti, Moldova, Pakistan, and Ukraine. Despite the generally non-significant coefficient for the production-type technology adoption (see Tables 6.4 and 6.6), this is the technolo-

Table 6.8: Countries statistically significant - Part 2

Countries	modern ICT	old ICT	transportation	production
Hungary	0.18(0.037)	-0.09(0.054)	0.11(0.015)	
Iceland	-0.33(0.003)		0.16(0.034)	
Indonesia	0.25(0.000)			
Iran		0.17(0.070)		
Ireland			0.16(0.000)	-0.08(0.078)
Israel				0.67(0.077)
Japan	-0.03(0.074)			
Jordan	-0.18(0.008)	0.31(0.068)		
Kenya	-0.19(0.043)			
Korea, Republic of			0.100(0.009)	
Kyrgyz Republic		0.33(0.046)		
Lao				0.90(0.026)
Latvia		-0.07(0.048)	0.133(0.028)	
Luxembourg				0.22(0.045)
Malawi		0.12(0.024)		
Malaysia			0.30(0.000)	
Mali		0.26(0.048)		
Mauritius				-0.28(0.000)
Mexico		0.12(0.002)		
Moldova		0.12(0.005)	-0.31(0.006)	0.25(0.000)
Morocco		-0.16(0.062)		
Netherlands	0.06(0.044)	0.09(0.008)		-0.12(0.051)
Norway				0.07(0.031)
Pakistan			-0.31(0.000)	-0.14(0.089)
Panama		0.27(0.016)		0.39 (0.018)

gy-type with the highest number of significant coefficients *per* country (34). However, the number of negatively significant coefficients (16) is relatively close to the number of positively significant coefficients (18). Countries with a significantly positive sign for the production technology-type coefficient are, e.g., Argentina, Bolivia, Israel, Poland, and Spain. Countries with a significantly negative sign for the production technology-type coefficient are, e.g., Bulgaria, Chile, Hong-Kong, Pakistan, and Portugal. This means that despite the fact that a positive effect of some types of technology adoption in raising inequality was obtained for the panel database, mainly for the rich countries (see Table 6.5), it is undoubted that we can identify both rich and poor countries with significantly positive signs and with significantly negative signs.

6.5 Conclusion

Quantitative assessments of the determinants of inequality are scarce. So are the quantitative evaluations of the theories that assume that skills are complementary to technologies and jointly determine the evolution of inequality. In this work we seek to contribute to enrich that literature. To this end, we use very recent data on the Gini index, available for a wide range of countries and years and relate it with measures of skill-complementary technological adoption of 20 different technologies. First, we analyze each skill-complementary technology and evaluate its effect on inequality. We discovered that several skill-complementary technologies contribute to the inequality rise and none contribute to the inequality drop. Adoption of technologies such as aviation, cell phones, electricity production, internet, telephone, and TV, contribute to increase inequality. Then, we construct four different measures of technol-

Table 6.9: Countries statistically significant - Part 3

Countries	modern ICT	old ICT	transportation	production
Paraguay				0.11(0.061)
Peru	0.30(0.012)			
Poland				0.27(0.019)
Portugal				-0.21(0.097)
Romania				-0.31(0.000)
Russian Federation		0.15(0.014)		-0.11(0.032)
Sierra Leone			0.21(0.000)	
Singapore		0.07(0.093)		
South Africa			0.08(0.086)	-0.26(0.000)
Spain				0.19(0.014)
Sri Lanka			0.21(0.001)	
Sweden		0.03(0.087)		
Switzerland	-0.13(0.002)			0.05(0.040)
Taiwan				0.14(0.012)
Tajikistan	0.13(0.062)	0.14(0.000)		
Tanzania			0.42(0.001)	
Thailand	0.17(0.006)	0.37(0.002)	0.21(0.004)	-0.34(0.003)
Uganda		0.72(0.000)		
Ukraine		0.15(0.020)	-0.510(0.000)	
United Kingdom	0.05(0.010)			
Uruguay	0.16(0.046)		0.32(0.001)	
Venezuela		0.19(0.002)	0.09(0.011)	-0.23(0.000)
Yugoslavia		1.49(0.032)		-1.59(0.017)
Zambia				0.87(0.002)

ogy-types allowing us to evaluate the conditional contribution of each type to the evolution of inequality. We found strong evidence that older ICT and transport technologies (and less frequently modern ICT) tend to increase inequality. Thus, earlier emphasis in the literature on the effect of ICT in raising inequality is relatively shaken by our results, as modern ICT adoption is definitively not the most significant type of technology adoption in raising inequality. We also discovered that results are much stronger in rich countries than in poor. The use of heterogenous panel estimators allowed us to highlight the diversity of results among countries. Nevertheless, an overwhelming number of countries present an influence of increased skill-complementary technological adoption (mainly in older ICT and transportation technologies) on the increase of inequality.

Our results are robust to a series of modifications in specification, estimator, samples, and to the skill-complementary technological adoption measure.

These results may have policy implications for the design of incentives for adopting technologies, especially for rich and well human-capital-endowed countries, in which the effect in the rise of inequality may be quite significant.

Chapter 7

Conclusions

In the first essay, we investigate the relationship between human capital and the ancestral genetic diversity of populations. The paper highlights a new channel through which genetic diversity can affect development, through human capital. We have devised a very simple model in which human capital benefits from an increasing (inherited) variety of genetic traits (heterozygosity), which enhance learning abilities. Additionally, a cost of human capital which depends increasingly on genetic diversity is essential to depict a hump-shaped relationship between genetic diversity and the human capital supply. This cost represents the additional effort economic agents have to support in order to overcome the negative influence of very diverse genetic backgrounds on the school environment. Despite its simplicity, the model encompasses the interplay between nature and nurture in the human capital supply of the economy, the presence of inherited family genetic traits and the costs of diversity on learning environment.

We based our empirical study on a database of human capital variables coming from Cohen and Soto (2007) - for measures of quantity of human capital and from Hanushek and Woessmann (2012) - for measures of quality of human capital and then merged it with the database of genetic diversity, from Ashraf and Galor (2013). We found a hump-shaped relationship between human capital and genetic diversity, confirming the idea that the influence of genetic diversity on development may be through human capital. A 1% change in low levels of genetic diversity may imply large effects in schooling that can oscillate between more 4 and 12[1] months of schooling (more 0.31 to 0.48 points in international tests scores) and negative effects when there is high genetic diversity (less 2 to 10 months of schooling and near less 0.3 to 0.4 points on scores).

We show and discuss a number of robustness tests with instrumental variables regressions. The overall conclusion is that the hump-shaped relationship between human capital and genetic diversity can indeed be regarded as a causal relationship. Thus, human capital outcomes may have been set onto their current paths millenia ago, when great human migrations shaped the countries' genetic diversity that we see today.

As noted before, Ashraf and Galor's (2013) have studied the effect of genetic diversity on development using technology as the theoretical link between genetic diversity and development. That article argued that an increase in diversity enhances production possibilities as a wider spectrum of traits is more likely to contain those that are complementary to the advancement of superior technologies. However, those authors used the population density in 1500 as a proxy for technological development. Alternatively, in the second essay of this thesis, we use direct measures of technological adoption provided by Comin *et. al.* (2010) to address such relationship. This study highlights a strong hump-shaped relationship between genetic diversity and technological developments in 1500. This means that some of the technological achievements may stem from the genetic diversity mostly determined more than a millennium ago. Results are robust to the introduction of several controls, and to IV estimation.

[1]Using the most preferable specifications (Tables 2.4 and 2.5 and regressions with controls and continent dummies).

In the third essay we collected the invention dates for more than 100 inventions from the seventeenth century around the world. With that we studied the determinants of the probability of the innovations occur in a given country, trying to contribute to the literature that explains the triggers of the industrial revolutions. We found evidence according to which the scale of the country, measured by its population, is an important determinant of the probability to innovate. This strong effect means that nearly more 30 inhabitants in a country could increase the probability to innovate from 12% to 41%. Our results show a small negative effect of education, reflecting the relatively lower importance of education as a source of innovations. This is according to the opinion of some economic historians which argue that the advent of formal education was posterior to the rise of innovations during the industrial revolution. Openness is rarely a significant determinant of inventions and when it is, it appears with a negative sign. Innovations from 1600 onwards are also related to previous technological knowledge in the country. We found a negative effect if the USA is included in the sample and a positive effect otherwise. This reflects the fact that the United States developed a strong industrial revolution without any substantial technological development in 1500. A standard-deviation increase in technology in 1500 (0.3) would increase the probability to innovate from 3% to 9%, when a positive effect occurs. We also found evidence for the influence of geographic distance and genetic distance on the probability of inventions. Generally and interestingly, distance to the UK and proximity to the USA increased the historical probability of innovation for a given country. In particular, geographical distance statistical significance is robust to all specification changes we have performed. However, the quantitative effects are small. An increased distance of 100 kilometers to the UK imply 0.2%-0.3% more probability of innovate. On the contrary, more 100km distance to the USA decreases the probability of innovation between 0.4% to 0.5%. More 100 in genetic distance to the UK (roughly 1/3 of the genetic distance between the USA and the UK) imply more 7%-9% more probability of innovating. Finally, we have tested the influence of the same regressors on a variable that intends to measure how early innovations occurred. Now, a rise in population by 30 persons would have deterred innovations by 47 to 51 years. More technology in 1500, say more 0.1, foster innovations in 14 to 17 years. Distance to UK and proximity to USA (both physical and genetic) did foster innovations. While the quantitative effect of physical distance is rather modest (a 100 km distance from the UK foster innovations in 1 year while 100 km closer to the USA foster innovations in 1 year and 1/2), the quantitative effect of genetic distance is more important: 100 genetic distance from the UK would have deterred innovations for more 22 years while 100 genetic proximity to the USA would have fostered innovation in 25 years.

There is scarce quantitative literature on the determinants of inequality. The last two essays of this thesis contribute to fill this gap on the literature. We contribute to that literature by evaluating potential determinants of inequality in a large panel of countries. Earlier attempts have faced problems with the coverage and quality of the income inequality data. We use a recent standardized measure of the Gini coefficient, due to Solt (2009) to evaluate human capital, TFP and openness as possible determinants of inequality. We conclude that this measure also needs to be corrected for differences in original data precision. Failure to do so would determine crucially different and misleading results concerning the influence of human capital on inequality. Fortunately, Solt (2009) also provides the means to implement such correction. We found that inequality data, as well as other macroeconomic variables, are subject to cross-country dependence and nonstationarity and so, newly developed econometric methods designed to analyze moderate T, moderate N panels should be employed (Eberhardt and Teal, 2010). We proceeded along this line and implemented cointegration tests to evaluate the (Granger) causal-

ity between human capital and inequality. Results indicate a strong channel from human capital to inequality. Regressions based on heterogeneous panels methods indicate that there is great heterogeneity concerning the effects of TFP and openness on inequality. Additionally, we found a positive statistically significant effect of human capital on inequality once the Gini coefficient is corrected for differences in its precision. This result is robust to several specification changes and measurement changes both in the inequality variable and in the human capital variable. Notably, the positive effect of human capital on inequality remains highly significant in methods robust to reverse causality. This does not dismiss that some heterogeneous effects between different countries are also present. Contrary to what may have been the current wisdom until now, it is not only tertiary education that tends to cause higher inequality, but the effect is highlighted with a measure that strongly weights lower levels of education, suggesting further research on the effect of primary education on inequality. These results suggest that theories that are not based on country heterogeneity to explain the relationship between technology, openness, and inequality may be unrealistic. In fact, institutions and history may be behind the heterogeneous effects of human capital, technology, and openness on inequality detected. Additionally, contrary to earlier evidence, the results in this paper suggest that human capital may be seen as the most important worldwide determinant of inequality.
As pointed out before, quantitative assessments of the determinants of inequality are scarce. So are the quantitative evaluations of the theories that assume that skills are complementary to technologies and jointly determine the evolution of inequality. In the last essay of this thesis we seek to contribute to enrich that literature. To this end, we use very recent data on the Gini index, available for a wide range of countries and years and relate it with measures of skill-complementary technological adoption of 20 different technologies. First, we analyze each skill-complementary technology and evaluate its effect on inequality. We discovered that several skill-complementary technologies contribute to the inequality rise and none contribute to the inequality drop. Adoption of technologies such as aviation, cell phones, electricity production, internet, telephone, and TV, contribute to increase inequality. Then, we construct four different measures of technology-types allowing us to evaluate the conditional contribution of each type to the evolution of inequality. We found strong evidence that older ICT and transport technologies (and less frequently modern ICT) tend to increase inequality. Thus, earlier emphasis in the literature on the effect of ICT in raising inequality is relatively shaken by our results, as modern ICT adoption is definitively not the most significant type of technology adoption in raising inequality. We also discovered that results are much stronger in rich countries than in poor. The use of heterogeneous panel estimators allowed us to highlight the diversity of results among countries. Nevertheless, an overwhelming number of countries present an influence of increased skill-complementary technological adoption (mainly in older ICT and transportation technologies) on the increase of inequality. Our results are robust to a series of modifications in specification, estimator, samples, and to the skill-complementary technological adoption measure. These results may have policy implications for the design of incentives for adopting technologies, especially for rich and well human-capital-endowed countries, in which the effect in the rise of inequality may be quite significant.

Bibliography

[AA11] S. Adobo and J. Agbayewa. Effect of Homogeneous and Heterogeneous Ability Grouping Class Teaching on Student's Interest, Attitude and Achievement in Integrated Science. *International Journal of Psychology and Counselling*, 3(3):48–54, 2011.

[ABDS11] T. Andersen, J. Bentzen, C-J. Dalgaard, and P. Sharp. Religious Orders and Growth Through Cultural Change in Pre-Industrial England. 2011.

[Ace2a] D. Acemoglu. Technical Change, Inequality, and the Labor Market. *Journal of Economic Literature*, 40(1):7–72, 2002a.

[Ace2b] D. Acemoglu. Directed Technical Change. *Review of Economic Studies*, 69:781–810, 2002b.

[Ace98] D. Acemoglu. Why do new technologies complement skills? directed technical change and wage inequality. *The Quarterly Journal of Economics, MIT Press*, 113(4):1055–1089, 1998.

[Ace03] D. Acemoglu. Technology and Inequality. *NBER Reporter: Winter 2003*, 2003.

[ADE+03] A. Alesina, A. Devleeschauwer, W. Easterly, S. Kurlat, and R. Wacziarg. Fractionalization. *Journal of Economic Growth*, 8(2):155–194, 2003.

[AG1a] Q. Ashraf and O. Galor. Dynamics and Stagnation in the Malthusian Epoch. *American Economic Review*, 101(August):2003–2041, 2011a.

[AG1b] Q. Ashraf and O. Galor. Cultural Diversity, Geographical Isolation, and the Origin of the Wealth of Nations. *NBER Working Papers 17640, National Bureau of Economic Research, Inc*, 2011b.

[AG13] Q. Ashraf and O. Galor. The Out of Africa Hypothesis, Human Genetic Diversity, and Comparative Economic Development. *American Economic Review*, 103(1):1–46, 2013.

[AH92] P. Aghion and P. Howitt. A Model of Growth Through Creative Destruction. *Econometrica*, 60:323–351, 1992.

[AHR12] A. Alesina, J. Harnoss, and H. Rapoport. *Birthplace Diversity and Economic Growth*. Manuscript, 2012.

[AHV02] P. Aghion, P. Howitt, and G. Violante. General Purpose Technology and Wage Inequality. *Journal of Economic Growth*, 7(4):315–345, 2002.

[AJR01] D. Acemoglu, S. Johnson, and J. Robinson. The Colonial Origins of Comparative Development: an Empirical Investigation. *American Economic Association*, 91(5):1369–1401, 2001.

[AJR05] D. Acemoglu, S. Johnson, and J. Robinson. Institutions as a Fundamental Cause of Long-Run Growth. *Handbook of Economic Growth*, 1(05):385–472, 2005.

[All9a] R. Allen. The British Industrial Revolution in Global Perspective. *Cambridge University Press, Cambridge*, 2009a.

[All9b] R. Allen. Engels' Pause: Technical Change, Capital Accumulation, and Inequality in the British Industrial Revolution. *Explorations in Economic History*, 46(4):418–435, 2009b.

[All03] R. Allen. Progress and Poverty in Early Modern Europe. *The Economic History Review New Series*, 56(3):403–443, 2003.

[ALPR11] L. Akinbami, X. Liu, P. Pastor, and C. Reuben. Attention Deficit Hyperactivity Disorder Among Children Aged 5–17 Years in the United States, 1998–2009. *NCHS Data Brief*, 70:August, 2011.

[AO12] P. Ahlerup and O. Olsson. The Roots of Ethnic Diversity. *Journal of Economic Growth*, 17(2):71–102, 2012.

[AOY09] P. Ahlerup, O. Olsson, and D. Yanagizawa. Social Capital vs Institutions in the Growth Process. *European Journal of Political Economy*, 25:1–14, 2009.

[AS12] G. Agostino and M. Scarlato. Inclusive Institutions, Innovation and Economic Growth: Estimates for European Countries. *MPRA Paper 43098, University Library of Munich, Germany*, 2012.

[Bak08] M. Baker. A Structural Model of the Transition to Agriculture. *Journal of Economic Growth*, 13(4):257–292, 2008.

[Bar91] R. Barro. Economic Growth in a Cross-Section of Countries. *The Quarterly Journal of Economics*, 106:407–443, 1991.

[Bar00] R. Barro. Inequality and Growth in a Panel of Countries. *Journal of Economic Growth*, 5(1):5–32, 2000.

[BCK+11] S. Broadberry, B. Campbell, A. Klein, M. Overton, and B. van Leeuwen. British Economic Growth, 1270-1870: An Output-based Approach. *University of Kent School of Economics Discussion Papers, KDPE 1203*, 2011.

[Bir01] J. Birchenall. Income Distribution , Human Capital and Economic Growth in Colombia. *Journal of Development Economics*, 66(1):271–287, 2001.

[BL01] R. Barro and J. Lee. International Data on Educational Attainment: Updates and Implications. *Oxford Economic Papers, Oxford University Press*, 53(3):541–63, 2001.

[BL13] R. Barro and J. Lee. A New Data Set of Educational Attainment in the World, 1950-2010. *Journal of Development Economics*, 104:184–198, 2013.

[BM03] R. Barro and R. McCleary. Religion and Economic Growth across Countries. *American Sociological Review*, 68(5):760–781, 2003.

[BvZ13] J. Bolt and J. van Zanden. The First Update of the Maddison Project; Re-Estimating Growth Before 1820. *Maddison Project Working Paper 4*, 2013.

[Car03] P. Carter. "Black" Cultural Capital, Status Positioning, and Schooling Conflicts for Low Income African American Youth. *Social Problems*, 50(1):136–155, 2003.

[Cas99] F. Caselli. Technological Revolutions. *American Economic Review*, 89:78–102, 1999.

[CC13] F. Caselli and W. Coleman. On the Theory of Ethnic Conflict. *Journal of the European Economic Association*, 11(S1):161–192, 2013.

[CEG10] D. Comin, W. Easterly, and E. Gong. Was the Wealth of Nations Determined in 1000 BC? *American Economic Journal: Macroeconomics*, 2(3):65–97, 2010.

[CH92] N. Crafts and C. Harley. Output Growth and the British Industrial Revolution: A Re-statement of the Crafts-Harley View. *Economic History Review*, 45(4):703–730, 1992.

[CH09] D. Comin and B. Hobijn. The CHAT Dataset. *NBER Working Papers 15319, National Bureau of Economic Research, Inc*, 2009.

[Cha00] A. Chakrabarti. Does Trade Cause Inequality? *Journal of Economic Development*, 25(2):1–21, 2000.

[CHR06] D. Comin, B. Hobijn, and E. Rovito. Five facts you need to know about technology diffusion. *NBER Working Papers 11928, National Bureau of Economic Research, Inc*, 2006.

[Cla96] G. Clark. The Political Foundations of Modern Economic Growth: England 1540-1800. *Journal of Interdisciplinary History*, 26(4):563–588, 1996.

[CM13] D. Comin and M. Mestieri. If technology has arrived everywhere, why has income diverged? pages NBER Working Papers 19010, National Bureau of Economic Research, Inc, 2013.

[CP13] A. Chudik and M. Pesaran. Common Correlated Effects Estimation of Heterogeneous Dynamic Panel Data Models with Weakly Exogenous Regressors. *Federal Reserve Bank of Dallas Globalization and Monetary Policy Institute Working Paper No.146*, 2013.

[Cra04] N. Crafts. Productivity Growth in the Industrial Revolution: A New Growth Accounting Perspective. *Journal of Economic History*, 64(2):521–535, 2004.

[CRHD13] D. Comin, E. Rossi-Hansberg, and M. Dmitriev. The Spatial Diffusion of Technology. *INET Research Notes 27, Institute for New Economic Thinking (INET)*, 2013.

[CS07] D. Cohen and M. Soto. Growth and Human Capital: Good Data, Good Results. *Journal of Economic Growth*, 12(1):51–76, 2007.

[DMR+11] S. Ding, L. Meriluoto, W. Reed, D. Tao, and H. Wu. The Impact of Agricultural Technology Adoption on Income Inequality in Rural China: Evidence From Southern Yunnan Province. *China Economic Review*, 22(3):344–356, 2011.

[DNMD+13] J-E. De Neve, S. Mikkaylov, C. Dawes, N. Cristakis, and J. Fowler. Born to Lead? A Twin Design and Genetic Association Study of Leadership Role Occupancy. *The Leadership Quarterly*, 24:45–60, 2013.

[DO13] C-J. Dalgaard and O. Olsson. Why Are Rich Countries More Politically Cohesive? *Scandinavian Journal of Economics*, 115(2):423–448, 2013.

[DS96] K. Deininger and L. Squire. New Data Set Measuring Income Inequality. *World Bank Economic Review*, 10:565–591, 1996.

[Eas81] R. Easterlin. Why Isn't the Whole World Developed? *Journal of Economic History*, 41(1):1–19, 1981.

[EB09] M. Eberhardt and S. Bond. Cross-Section Dependence in Nonstationary Panel Models: a Novel Estimator. *MPRA Paper No. 17870*, 2009.

[EHS13] M. Eberhardt, C. Helmers, and H. Strauss. Do Spillovers Matter When Estimating Private Returns to R&D? *The Review of Economics and Statistics*, 95(2):436–448, 2013.

[EL97] W. Easterly and R. Levine. Africa's Growth Tragedy: Policies and Ethnic Divisions. *Quarterly Journal of Economics*, 112(4):1203–1250, 1997.

[EL12] W. Easterly and R. Levine. The European Origins of Economic Development. *NBER Working Paper No. 18162, National Bureau of Economic Research, Inc*, 2012.

[EP14] M. Eberhardt and A. Presbitero. This Time They Are Different: Heterogeneity and Nonlinearity in the Relationship Between Debt and Growth. *version 13th February 2014*, 2014.

[ET0a] M. Eberhardt and F. Teal. Aggregation versus Heterogeneity in Cross-Country Growth Empirics. *Economics Series Working Papers CSAE WPS/2010-32, University of Oxford, Department of Economics.*, 2010a.

[ET0b] M. Eberhardt and F. Teal. Productivity Analysis in Global Manufacturing Production. *Department of Economics Discussion Paper Series (Ref: 515), University of Oxford, Department of Economics.*, 2010b.

[ET3a] M. Eberhardt and F. Teal. No Man goes in the Tundra: Spatial Heterogeneity in Agricultural Productivity Analysis. *Oxford Bulletin of Economics and Statistics*, 75(6):914–939, 2013a.

[ET3b] M. Eberhardt and F. Teal. Structural Change and Cross-Country Growth Empirics. *World Bank Economic Review*, 27(2):229–271, 2013b.

[ET11] M. Eberhardt and F. Teal. Econometrics for Grumblers: a New Look At the Literature on Cross-Country Growth Empirics. *Journal of Economic Surveys*, 25(1):109–155, 2011.

[FB09] C. Ferguson and K. Beaver. Natural Born Killers: The Genetic Origins of Extreme Violence. *Aggression and Violent Behavior*, 14:285–294, 2009.

[FIT08] R. Feenstra, R. Inklaar, and M. Timmer. The Next Generation of the Penn World Table. *NBER Working Papers 19255*, National Bureau of Economic Research, Inc, 2008.

[Gal05] O. Galor. From Stagnation to Growth: Unified Growth Theory. *Handbook of Economic Growth*, Vol. 1A, ch. 4:Edited by Philippe Aghion and Steven N. Durlauf, Amsterdam, The Netherlands, Elsevier, North-Holland, 2005.

[GLPLdSS04] E. Glaeser, R. La Porta, F. Lopez-de Silanes, and A. Shleifer. Do Institutions Cause Growth? *Journal of Economic Growth*, 9(3):271–303, 2004.

[GM00] O. Galor and O. Moav. Ability-Biased Technological Transition, Wage Inequality, and Economic Growth. *The Quarterly Journal of Economics*, 115(2):469–497, 2000.

[GM06] O. Galor and O. Moav. Das Human-Kapital: A Theory of the Demise of the Class Structure. *Review of Economic Studies*, 73(1):85–117, 2006.

[GO07] D. Greasley and L. Oxley. Patenting, Intellectual Property Rights and Sectoral Outputs in Industrial Revolution Britain, 1780-1851. *Journal of Econometrics*, 139:340–354, 2007.

[GS01] J. Gallup and J. Sachs. The Economic Burden of Malaria. *American Journal of Tropical Medicine and Hygiene*, 64(1-2):85–96, 2001.

[GS12] M. Gómez and T. Sequeira. Phases of Economic Development: Do Initial Endowments Matter? *The BE Journal of Macroeconomics (Topics)*, 12(1):Article 21, 2012.

[GSZ09] L. Guiso, P. Sapienza, and L. Zingales. Cultural Biases in Economic Exchange? *Quarterly Journal of Economics*, 124(3):1095–1131, 2009.

[GT97] O. Galor and D. Tsiddon. Technological Progress, Mobility, and Economic Growth. *American Economic Review*, 87:362–382, 1997.

[GT09] C. Gutiérrez and R. Tanaka. Inequality and Education Decisions in Developing Countries. *The Journal of Economic Inequality*, 7(1):55–81, 2009.

[GY97] J. Greenwood and M. Yorukoglu. 1974. *Carnegie-Rochester Conference Series on Public Policy 46*, pages 49–96, 1997.

[HDL+14] W. Hill, G. Davies, L. Lagemaat, A. Christoforou, R. Marioni, C. Fernandes, D. Liewald, M. Croning, A. Payton, L. Craig, L. Whalley, M. Horan, W. Ollier, N. Hansell, M. Wright, N. Martin, G. Montgomery, V. Steen, S. Hellard, T. Espeseth, A. Lundervold, I. Reinvang, J. Starr, N. Pendleton, S. Grant, T. Bates, and I. Deary. Human Cognitive Ability is Influenced by Genetic Variation in Components of Post-synaptic Signalling Complexes Assembled by NMDA Receptors and MAGUK Proteins. *Translational Psichiatry*, 4:1–8, 2014.

[HJ99] R. Hall and C. Jones. Why Do Some Countries Produce So Much More Output Per Worker Than Others? *Quarterly Journal of Economics*, 114(1):83–116, 1999.

[HKV05] A. Hornstein, P. Krusell, and G. Violante. Chapter 20 The Effects of Technical Change on Labor Market Inequalities. *Handbook of Economic Growth*, 1:1275–1370, 2005.

[Hou02] R. Houston. Literacy in Early Modern Europe - Culture and Education 1500-1800. *Second Edition. Pearson Education*, 2002.

[HW11] E. Hanushek and L. Woessmann. The Economics of International Differences in Educational Achievement (chapter 2). *Handbook of the Economics of Education*, 3:89–200, 2011.

[HW12] E. Hanushek and L. Woessmann. Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation. *Journal of Economic Growth*, 17(4):267–321, 2012.

[Iac10] M. Iacopetta. Phases of Economic Development and the Transitional Dynamics of an Innovation-education Growth Model. *European Economic Review*, 54(2):317–330, 2010.

[Iyi11] M. Iyigun. On the Socioeconomic Consequences of Religious Strife and Coexistence. *The Oxford handbook of the economics of religion*, pages 217–231, 2011.

[JLP13] F. Jaumotte, S. Lall, and C. Papageorgiou. Rising Income Inequality: Technology, or Trade and Financial Globalization? *IMF Economic Review*, 61(2):271–309, 2013.

[JNSS09] D. Johnston, M. Nicholls, M. Shah, and M. Shields. Nature's Experiment? Handedness and Early Childhood Development. *Demography*, 46(2):281–301, 2009.

[Kor14] T. Koriath. Autism Prevalence Now 1 in 68, Varies by Sex, Race/Ethnic Group. AAP News - Breaking News, March 27. 2014.

[Kuz55] S. Kuznets. Economic Growth and Income Inequality. *American Economic Review*, 45(1):1–28, 1955.

[Lad13] A. Lada. Clash of Brothers in a Contagious World : Wars to Avoid Diffusion. *Institute of Economics, Centre for Economic and Regional Studies, Hungarian Academy of Sciences, IEHAS Discussion Papers, 1333*, 2013.

[Lin99] J. Lin. Technological Change and Agricultural Household Income Distribution : Theory and Evidence From. *The Australian Journal of Agriculture and Resource Economics*, 43(2):179–194, 1999.

[LP99] R. La Porta. The Quality of Government. *Journal of Law, Economics, and Organization*, 15(1):222–279, 1999.

[LS08] I. Lindner and H. Strulik. Social Fractionalization, Endogenous Appropriation Norms, and Economic Development. *Economica*, 75(298):244–258, 2008.

[Luc88] Jr. Lucas. On the Mecanics of Economic Development. *Journal of Monetary Economics*, 22:3–42, 1988.

[McM09] I. McManus. The History and Geography of Human Handedness. *Language Lateralization and Psychosis*, pages ed. Iris E. C. Sommer and René S. Kahn., Cambridge University Press, 2009.

[Mit93] D. Mitch. The Role of Human Capital in the First Industrial Revolution. *The British Industrial Revolution: An Economic Perspective*, pages 267–307, 1993.

[Mit98] B. Mitchell. Historical Statistics. *4th Edition, New York: Stockton Press*, 1998.

[Mok90] J. Mokyr. The Lever of Riches. *Oxford University Press*, 1990.

[Mok05] J. Mokyr. Mobility, Creativity and Technological Development. *Prepared for the session on "Creativity and the Economy", German Association of Philosophy, Berlin, Sept. 18*, 2005.

[Mok09] J. Mokyr. The Enlightened Economy: An Economic History of Britain 1700-1850. *New Haven: Yale University Press*, 2009.

[MP04] P. Martins and P. Pereira. Does Education Reduce Wage Inequality? Quantile Regression Evidence From 16 Countries. *Labour Economics*, 11:355–371, 2004.

[MRW92] N. Mankiw, D. Romer, and D. Weil. A Contribution to the Empirics of Economic Growth. *The Quarterly Journal of Economics*, 107(2):407–437, 1992.

[MZ11] T. Mayer and S. Zignago. Notes on CEPIIs distances measures : the GeoDist Database. *CEPII Working Paper 2011-25*, 2011.

[NS14] P. Neves and S. Silva. Inequality and Growth: Uncovering the Main Conclusions from the Empirics. *Journal of Development Studies*, 50(1):1–21, 2014.

[OH05] O. Olsson and Jr. Hibbs. Biogeography and Long-Run Economic Development. *European Economic Review*, 49:909–938, 2005.

[Ols07] O. Olsson. Conflict Diamonds. *Journal of Development Economics*, 82:267–286, 2007.

[PC01] R. Plomin and I. Craig. Genetics, Environment and Cognitive Abilities: Review and Work in Progress Towards a Genome Scan for Quantitative Trait Locus Associations Using DNA Pooling. *British Journal of Psychiatry*, 178:s41–s48, 2001.

[Pes04] M. Pesaran. General Diagnostic Tests for Cross Section Dependence in Panels. *IZA Discussion Papers 1240*, 2004.

[Pes06] M. Pesaran. Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure. *Econometrica*, 74(4):967–1012, 2006.

[Pes07] M. Pesaran. A Simple Panel Unit Root Test in the Presence of Cross-Section Dependence. *Journal of Applied Econometrics*, 22(2):265–312, 2007.

[PH10] I. Pyysiäinen and M. Hauser. The Origins of Religion : Evolved Adaptation or By-Product? *Trends in Cognitive Sciences*, 14(3):104–109, 2010.

[PLH+04] S. Petril, P. Lipton, J. Hewitt, R. Plomin, S. Cherny, R. Corley, and J. DeFries. Genetic and Environmental Contributions to General Cognitive Ability Through the First 16 Years of Life. *Developmental Psichology*, 40(5):805–812, 2004.

[PM00] P. Phillips and H. Moon. Nonstationary Panel Data Analysis: An Overview of Some Recent Developments. *Econometric Reviews*, 19(3):263–286, 2000.

[PS12] L. Ponciano and A. Shabazian. Interculturalism: Addressing Diversity in Early Childhood. Dimensions of Early Childhood. 40(1):23–29, 2012.

[Psa94] G. Psacharopoulos. Returns to Investment in Education:A Global Update. *World Development*, 22(9):1325–1343, 1994.

[PSY13] M. Pesaran, V. Smith, and T. Yamagata. Panel Unit Root Tests in the Presence of a Multifactor Error Structure. *Journal of Econometrics*, 175(2):94–115, 2013.

[PT11] M. Pesaran and E. Tosetti. Large Panels with Common Factors and Spatial Correlations. *Journal of Econometrics*, 161(2):182–202, 2011.

[Reb91] S. Rebelo. Long-Run Policy Analysis and Long-Run Growth. *University of Chicago Press*, 99(3):500–521, 1991.

[Ric95] J. Richardson. Income Inequality and Trade: How to Think, What to Conclude. *Journal of Economic Perspectives*, 9(3):33–55, 1995.

[RMD+13] C. Rietveld, S. Medland, J. Derringer, J. Yang, T. Esko, N. Martin, Harm-Jan Westra, K. Shakhbazov, A.l Abdellaoui, A. Agrawal, et al. Gwas of 126,559 individuals identifies genetic variants associated with educational attainment. *Science*, 340(6139):1467–1471, 2013.

[Rom90] P. Romer. Endogenous Technological Change. *Journal of Political Economy*, 98(5):October, part II, S71–S102, 1990.

[RS13] J. Rattsø and H. Stokke. Regional Convergence of Income and Education: Investigation of Distribution Dynamics. *Urban Studies*, 2013.

[RST04] D. Rodrik, A. Subramanian, and F. Trebbi. Institutions Rule: The Primacy of Institutions Over Geography and Integration in Economic Development. *Journal of Economic Growth*, 9(2):131–165, 2004.

[RTZ3a] D. Rohner, M. Thoenig, and F. Zilibotti. Seeds of distrust: Conflict in Uganda. *Journal of Economic Growth*, 18(3):217–252, 2013a.

[RTZ3b] D. Rohner, M. Thoenig, and F. Zilibotti. War Signals : A Theory of Trade, Trust and Conflict War Signals : A Theory of Trade , Trust and Conflict. *Review of Economic Studies, Oxford University Press*, 80(3):1114–1147, 2013b.

[SBHS11] V. Stocké, H. Blossfeld, K. Hoenig, and M. Sixt. 7 Social Inequality and Educational Decisions in the Life Course. *Zeitschrift für Erziehungswissenschaft*, 14(2):103–119, 2011.

[SE00] K. Sokoloff and S. Engerman. History Lessons: Institutions, Factors Endowments, and Paths of Development in the New World. *The Journal of Economic Perspectives*, 14(3):217–232, 2000.

[Seq09] T. Sequeira. On the Relationship Between Education and Country-Risk: A New Fact. *Acta Oeconomica*, 59(4):441–454, 2009.

[SF09] T. Sequeira and N. Ferraz. Is Education Prejudiced by Country-Risk? *Economic Record*, 85(270):307–317, 2009.

[SKC+14] M-H. Shin, S-S. Kweon, J-S. Choi, Y-H. Lee, H-S. Nam, K-S. Park, H. Kim, H-R. Song, B. Kim, S-M. Choi, S-Y. Oh, and S-K Jeong. The Effect of an APOE Polymorphism on Cognitive Function Depends on Age. *Journal of Neurology*, 261:66–72, 2014.

[Sol09] F. Solt. Standardizing the World Income Inequality Database. *Social Science Quarterly*, 90(2):231–242, 2009.

[Sto01] N. Stockey. A Quantitative Model of the British Industrial Revolution, 1780-1850. *Carnegie-Rochester Conference Series on Public Policy*, 55:55–109, 2001.

[SW3a] E. Spolaore and R. Wacziarg. How Deep Are the Roots of Economic Development? *Journal of Economic Literature*, 51(2):325–369, 2013a.

[SW3b] E. Spolaore and R. Wacziarg. War and relatedness. *CAGE Online Working Paper Series 140, Competitive Advantage in the Global Economy (CAGE)*, 2013b.

[SW09] E. Spolaore and R Wacziarg. The Diffusion of Development. *The Quarterly Journal of Economics*, (May), 2009.

[SW12] E. Spolaore and R. Wacziarg. Long-Term Barriers to the International Diffusion of Inventions. *NBER International Seminar on Macroeconomics*, 8(1):11–46, 2012.

[Tab08] G. Tabellini. Institutions and Culture. *Journal of the European Economic Association*, 6(2-3):255–294, 2008.

[TDH12] E. Tucker-Drob and P. Harden. Learning Motivation Mediates Gene-by-Socioeconomic Status Interaction on Mathematics Achievement in Early Childhood. *Learning and Individual Differences*, 22:37–45, 2012.

[TI10] N. Terry and M. Irving. Cultural and Linguistic 4 Diversity: Issues In Education, In Special Education For All Teachers. *Ed. Colarusso and O'rourke*, 5th Ed:Kendall Hunt Publishing Co, 2010.

[Wan11] L. Wang. How Does Education Affect the Earnings Distribution in Urban China? *IZA Discussion Paper No. 6173*, 2011.

[Wes07] J. Westerlund. Testing for Error Correction in Panel Data. *Oxford Bulletin of Economics and Statistics*, 69(6):709-748, 2007.

Appendix A

Appendix of Chapter 2

A.1 List of countries - largest (123) sample

Afghanistan; United Arab Emirates; Argentina; Australia; Austria; Burundi; Belgium; Benin; Bangladesh; Bulgaria; Bahrain; Bolivia; Brazil; Botswana; Central African Republic; Canada; Switzerland; Chile; China; Cameroon; Congo, Rep.; Colombia; Costa Rica; Cuba; Cyprus; Czech Republic; Germany; Denmark; Dominican Republic; Algeria; Ecuador; Egypt, Arab Rep.; Spain; Estonia; Ethiopia; Finland; Fiji; France; United Kingdom; Ghana; Gambia, The; Greece; Guatemala; Guyana; Hong Kong; China; Honduras; Croatia; Haiti; Hungary; Indonesia; India; Ireland; Iran; Islamic Rep.; Iraq; Israel; Italy; Jamaica; Jordan; Japan; Kazakhstan; Kenya; Korea, Rep.; Kuwait; Liberia; Libya; Sri Lanka; Lesotho; Lithuania; Latvia; Moldova; Mexico; Mali; Malta; Myanmar; Mozambique; Mauritania; Mauritius; Malawi; Malaysia; Niger; Nicaragua; Netherlands; Norway; Nepal; New Zealand; Pakistan; Panama; Peru; Philippines; Papua New Guinea; Poland; Puerto Rico; Portugal; Paraguay; Romania; Russian Federation; Rwanda; Sudan; Senegal; Singapore; Sierra Leone; El Salvador; Slovak Republic; Slovenia; Sweden; Swaziland; Syrian Arab Republic; Togo; Thailand; Tajikistan; Trinidad and Tobago; Tunisia; Turkey; Uganda; Uruguay; United States; Venezuela, RB; Vietnam; Serbia and Montenegro; South Africa; Congo, Dem. Rep.; Zambia; Zimbabwe.

A.2 Alternative Specifications

In this subsection, we present results for alternative specifications, mentioned in the text.

Table A.2: Human Capital and Genetic Diversity (other controls)

	Mobility index-predicted genetic diversity	Mobility index-predicted genetic diversity square	Social infrastructure	% of population at risk of contracting malaria	$Adj.\ R^2$ / Observations
(1)	227.93**	-156.42**	5.82***	-2.53***	0.68
	(105.18)	(74.28)	(0.90)	(0.55)	78
(2)	248.92***	-168.62***	3.61***	-0.80***	0.68
	(86.44)	(61.35)	(0.58)	(0.25)	47
(3)	7.19	-4.70	0.25***	-0.10***	0.49
	(8.04)	(5.64)	(0.06)	(0.03)	72
(4)	89.17	-59.07	7.19***	-2.65***	0.72
	(133.93)	(95.45)	(0.86)	(0.56)	72
(5)	108.63	-71.69	7.78***	-2.75***	0.73
	(139.96)	(99.64)	(0.90)	(0.59)	72
(6)	284.70***	-200.08***	1.08***	-0.36	0.46
	(73.03)	(52.64)	(0.27)	(0.30)	44
(7)	279.69***	-196.47***	1.10***	-0.47	0.42
	(75.24)	(54.18)	(0.29)	(0.39)	44
(8)	113.78***	-79.84***	0.37***	-0.06	0.41
	(32.74)	(23.65)	(0.11)	(0.13)	44
(9)	16.02**	-11.26**	0.079***	-0.05***	0.44
	(6.27)	(4.48)	(0.02)	(0.02)	44
(10)	3856.87**	-2717.58**	28.06***	-23.38***	0.57
	(1598.96)	(1146.60)	(5.78)	(6.96)	40
(11)	3879.79**	-2733.94**	28.19***	-23.82***	0.57
	(1611.17)	(1154.79)	(5.83)	(6.88)	40
(12)	933.54***	-654.01***	6.08***	-4.03***	0.61
	(298.61)	(215.11)	(1.11)	(1.21)	40
(13)	116.71**	-81.54**	0.77***	-0.55**	0.50
	(51.31)	(36.55)	(0.16)	(0.21)	40

Note: Dependent Variables - (1) Years of schooling; (2) Years of schooling∗Interpersonal Trust; (3) % of population aged 15 or over with complete secondary education; (4) Years of schooling of population 15 and over, whether studying or not; (5) Years of schooling of population 15-64 who are not studying; (6) Average test score in math and science, primary through end of secondary school, all years (scaled to PISA scale divided by 100); (7) Average test score in math and science, only lower secondary, all years (scaled to PISA scale divided by 100); (8) Share of students reaching basic literacy (based on average test scores in math and science, primary through end of secondary school, all years); (9) Share of top-performing students (based on average test scores in math and science, primary through end of secondary school, all years); (10) Years of schooling∗Average test score in math and science, primary through end of secondary school; (11) Years of schooling∗Average test score in math and science, only lower secondary; (12) Years of schooling∗Share of students reaching basic literacy (based on average test scores in math and science, primary through end of secondary school; (13) Years of schooling∗Share of top-performing students (based on average test scores in math and science, primary through end of secondary school, all years). F-tests for the Predict Genetic Diversity Squared coefficient (not shown) always rejects for significant coefficients. Results are available upon request.

Level of significance: *** for p-value<0.01; **for p-value<0.05;* for p-value<0.1.

Values between parentheses are standard errors.

Appendix B

Appendix of Chapter 3

B.1 List of countries - largest (106) sample

Afghanistan; Angola; Argentina; Australia; Austria; Belgium; Benin; Burkina Faso; Bangladesh; Bosnia And Herzegovina; Belize; Bolivia; Brazil; Botswana; Central African Republic; Canada; Switzerland; Chile; China; CÃ´te D'Ivoire; Cameroon; Congo, Rep.; Colombia; Costa Rica; Cuba; Germany; Denmark; Algeria; Ecuador; Egypt, Arab Rep.; Spain; Ethiopia; Finland; France; Gabon; Ghana; Guinea; Guinea-Bissau; Greece; Guatemala; Guyana; Honduras; Hungary; Indonesia; India; Ireland; Iran, Islamic Rep.; Iraq; Italy; Japan; Kenya; Cambodia; Lao Pdr; Liberia; Libya; Lesotho; Lithuania; Morocco; Madagascar; Mexico; Mali; Myanmar; Mongolia; Mauritania; Malaysia; Namibia; Niger; Nigeria; Nicaragua; Netherlands; Norway; Nepal; New Zealand; Pakistan; Panama; Peru; Philippines; Papua New Guinea; Poland; Portugal; Paraguay; Romania; Russian Federation; Saudi Arabia; Sudan; Senegal; Sierra Leone; El Salvador; Sweden; Syrian Arab Republic; Chad; Thailand; Tunisia; Turkey; Tanzania; Uganda; Ukraine; Uruguay; United States; Uzbekistan; Venezuela; Vietnam; South Africa; Congo, Dem. Rep.; Zambia; Zimbabwe.

B.2 2SLS Regressions for Different Technologies

Table B.2.1: Technological Adoption and Genetic Diversity (2SLS Estimates)

	(1)	(2)	(3)	(4)	(5)
Dependent var.	*agr*	*agr*	*agr*	*agr*	*agr*
Observed diversity	19.08 (0.591)	-	-	-	-
Observed diversity Square	-16.70 (0.567)	-	-	-	-
Predicted Diversity	-	56.84*** (0.008)	55.78*** (0.009)	-	-
Predicted Diversity Square	-	-40.86*** (0.009)	-40.14** (0.010)	-	-
Mobility Index-Predicted Genetic Diversity	-	-	-	44.7*** (0.006)	43.39*** (0.008)
Mobility Index-Predicted Genetic Diversity Square	-	-	-	-31.84*** (0.009)	-30.83** (0.011)
log Neolithic Transition Time	.37** (0.013)	.24*** (0.000)	.24*** (0.000)	.34*** (0.000)	.34*** (0.000)
log Percentage Arable Land	0.05 (0.525)	.05* (0.086)	.05* (0.081)	0.03 (0.321)	0.03 (0.322)
log Absolute Latitude	0.02 (0.653)	-0.03 (0.274)	-0.03 (0.273)	-0.04 (0.181)	-0.04 (0.190)
log Land Suit. for Agriculture	-0.03 (0.599)	-0.01 (0.784)	-0.01 (0.775)	0.00 (0.893)	0.00 (0.902)
Kleibergen-Paap rk lm Statistic	-	41.2*** (0.000)	41.7*** (0.000)	30.278 (0.000)	29.469 (0.000)
Cragg-Donald Wald F Statistic	6.64^{δ}	$446.8^{\dagger\dagger\dagger}$ (11.04)	$353.7^{\dagger\dagger\dagger}$ (13.97)	151.996 (11.04)	145.476 (13.97)
Stock-Wright lm s Statistic	0.51 (0.775)	5.89 (0.207)	6.31 (0.277)	5.2 (0.268)	5.41 (0.368)
Hansen J-Statistic	exact. id.	1.88 (0.39)	2.48 (0.479)	1.03 (0.599)	1.60 (0.659)
Endog. Test	0.63 (0.427)	3.79 (0.151)	3.52 (0.172)	3.6 (0.166)	3.1 (0.212)
N	19	106	106	94	94

Note: Excluded instruments - (1) 2 instruments: first-stage fitted values of observed genetic diversity (square) and migratory distance from east Africa, equation is exactly identified; (2) and (4) 4 instruments: aerial distance from east Africa, aerial distance from east Africa (square), terrestrial distance from London and terrestrial distance from London (square); (3) and (5) 5 instruments: the same as in column (2) and (4) plus geodesic centroid latitude; Kleibergen-Paap rk lm statistic tests underidentification, under the null of the matrix of reduced form coefficients has rank=k1-1. Cragg-Donald Wald F statistic tests the null under which equation is weakly identified. this is compared with the Stock-Yogo weak id test critical values, which are reported in parentheses in that line. stock-wright lm s statistic tests the null under which the joint endogenous regressors have null coefficients. Hansen J-statistic tests the null under which the instruments are valid, i.e., uncorrelated with the error term.

Level of significance: *** for p-value<0.01; **for p-value<0.05;* for p-value<0.1. δ in column (1) 4.58 is the critical value for 15% critical iv size; † † † in columns (2) and (4), 11.04 is the critical value for relative IV bias of 5% (of the OLS bias); in columns (3) and (5), 13.97 is the critical value for relative IV bias of 5% (of the OLS bias).

values inside parentheses are p-values, except for the Cragg-Donald Wald F statistic. all regressions include continent dummies.

Table B.2.2: Technological Adoption and Genetic Diversity (2SLS Estimates)

	(1)	(2)	(3)	(4)	(5)
Dependent Var.	*com*	*com*	*com*	*com*	*com*
Observed Diversity	91.15*** (0.000)	-	-	-	-
Observed Diversity Square	-73.73*** (0.000)	-	-	-	-
Predicted Diversity	-	50.42*** (0.000)	49.42*** (0.000)	-	-
Predicted Diversity Square	-	-37.06*** (0.000)	-36.37*** (0.000)	-	-
Mobility Index-Predicted Genetic Diversity	-	-	-	36.42*** (0.000)	33.61*** (0.001)
Mobility Index-Predicted Genetic Diversity Square	-	-	-	-26.15*** (0.001)	-23.99*** (0.001)
log Neolithic Transition Time	.19* (0.071)	.23*** (0.000)	.23*** (0.000)	.28*** (0.000)	.28*** (0.000)
log Percentage Arable Land	.2*** (0.000)	.05* (0.053)	.05** (0.049)	0.04 (0.261)	0.04 (0.264)
log Absolute Latitude	.13*** (0.000)	.07*** (0.000)	.07*** (0.000)	.07*** (0.001)	.07*** (0.001)
log Land Suit. for Agriculture	0.02 (0.586)	0.00 (0.908)	0.00 (0.895)	0.00 (0.916)	0.00 (0.887)
Kleibergen-Paap rk lm Statistic	-	41.2*** (0.000)	41.7*** (0.000)	30.278 (0.000)	29.469 (0.000)
Cragg-Donald Wald F Statistic	6.64^{δ}	$446.8^{\dagger\dagger\dagger}$ (11.04)	$353.7^{\dagger\dagger\dagger}$ (13.97)	151.996 (11.04)	145.476 (13.97)
Stock-Wright lm s Statistic	5.97* (0.051)	11.43** (0.022)	11.55** (0.042)	12.79** (0.012)	14.29** (0.014)
Hansen J-Statistic	exact. id.	6.46** (0.04)	6.64* (0.084)	4.78* (0.092)	7.1* (0.069)
Endog. Test	1.43 (0.233)	1.44 (0.487)	2.34 (0.310)	3.59 (0.167)	2.16 (0.340)
N	19	106	106	94	94

Note: Excluded instruments - (1) 2 instruments: first-stage fitted values of observed genetic diversity (square) and migratory distance from east Africa, equation is exactly identified; (2) and (4) 4 instruments: aerial distance from east Africa, aerial distance from east Africa (square), terrestrial distance from London and terrestrial distance from London (square); (3) and (5) 5 instruments: the same as in column (2) and (4) plus geodesic centroid latitude; Kleibergen-Paap rk lm statistic tests underidentification, under the null of the matrix of reduced form coefficients has rank=k1-1. Cragg-Donald Wald F statistic tests the null under which equation is weakly identified. this is compared with the Stock-Yogo weak id test critical values, which are reported in parentheses in that line. stock-wright lm s statistic tests the null under which the joint endogenous regressors have null coefficients. Hansen J-statistic tests the null under which the instruments are valid, i.e., uncorrelated with the error term.

Level of significance: *** for p-value<0.01; **for p-value<0.05;* for p-value<0.1. δ in column (1) 4.58 is the critical value for 15% critical iv size; † † † in columns (2) and (4), 11.04 is the critical value for relative IV bias of 5% (of the OLS bias); in columns (3) and (5), 13.97 is the critical value for relative IV bias of 5% (of the OLS bias).

values inside parentheses are p-values, except for the Cragg-Donald Wald F statistic. all regressions include continent dummies.

Table B.2.3: Technological Adoption and Genetic Diversity (2SLS Estimates)

Dependent Var.	(1) *tra*	(2) *tra*	(3) *tra*	(4) *tra*	(5) *tra*
Observed Diversity	35.13 (0.174)	-	-	-	-
Observed Diversity Square	-29.88 (0.159)	-	-	-	-
Predicted Diversity	-	5.54 (0.483)	5.54 (0.482)	-	-
Predicted Diversity Square	-	-3.70 (0.550)	-3.70 (0.549)	-	-
Mobility Index-Predicted Genetic Diversity	-	-	-	5.29 (0.378)	5.35 (0.370)
Mobility Index-Predicted Genetic Diversity Square	-	-	-	-3.78 (0.428)	-3.83 (0.420)
log Neolithic Transition Time	.24** (0.024)	.18*** (0.000)	.18*** (0.000)	.23*** (0.000)	.23*** (0.000)
log Percentage Arable Land	.12** (0.046)	.04** (0.026)	.04** (0.026)	0.02 (0.251)	0.02 (0.251)
log Absolute Latitude	0.03 (0.331)	.04** (0.028)	.04** (0.028)	.04* (0.051)	.04* (0.052)
log Land Suit. for Agriculture	-.08** (0.040)	-.04* (0.055)	-.04* (0.055)	-0.03 (0.132)	-0.03 (0.133)
Kleibergen-Paap rk lm Statistic	-	41.2*** (0.000)	41.7*** (0.000)	30.278 (0.000)	29.469 (0.000)
Cragg-Donald Wald F Statistic	6.64^{δ}	$446.8^{\dagger\dagger\dagger}$ (11.04)	$353.7^{\dagger\dagger\dagger}$ (13.97)	151.996 (11.04)	145.476 (13.97)
Stock-Wright lm s Statistic	1.63 (0.442)	4.58 (0.333)	4.86 (0.433)	4.71 (0.318)	4.76 (0.446)
Hansen J-Statistic	exact. id.	2.50 (0.286)	2.57 (0.463)	1.50 (0.472)	1.51 (0.68)
Endog. Test	0.71 (0.399)	11.33*** (0.004)	12.47*** (0.002)	5.50* (0.064)	7.90** (0.019)
N	19	106	106	94	94

Note: Excluded instruments - (1) 2 instruments: first-stage fitted values of observed genetic diversity (square) and migratory distance from east Africa, equation is exactly identified; (2) and (4) 4 instruments: aerial distance from east Africa, aerial distance from east Africa (square), terrestrial distance from London and terrestrial distance from London (square); (3) and (5) 5 instruments: the same as in column (2) and (4) plus geodesic centroid latitude; Kleibergen-Paap rk lm statistic tests underidentification, under the null of the matrix of reduced form coefficients has rank=k1-1. Cragg-Donald Wald F statistic tests the null under which equation is weakly identified. this is compared with the Stock-Yogo weak id test critical values, which are reported in parentheses in that line. stock-wright lm s statistic tests the null under which the joint endogenous regressors have null coefficients. Hansen J-statistic tests the null under which the instruments are valid, i.e., uncorrelated with the error term.

Level of significance: *** for p-value<0.01; **for p-value<0.05;* for p-value<0.1. δ in column (1) 4.58 is the critical value for 15% critical iv size; † † † in columns (2) and (4), 11.04 is the critical value for relative IV bias of 5% (of the OLS bias); in columns (3) and (5), 13.97 is the critical value for relative IV bias of 5% (of the OLS bias).

values inside parentheses are p-values, except for the Cragg-Donald Wald F statistic. all regressions include continent dummies.

Table B.2.4: Technological Adoption and Genetic Diversity (2SLS Estimates)

Dependent Var.	(1) *mil*	(2) *mil*	(3) *mil*	(4) *mil*	(5) *mil*
Observed Diversity	44.70 (0.128)	-	-	-	-
Observed Diversity Square	-37.50 (0.120)	-	-	-	-
Predicted Diversity	-	8.02 (0.501)	7.89 (0.510)	-	-
Predicted Diversity Square	-	-6.28 (0.505)	-6.19 (0.513)	-	-
Mobility Index-Predicted Genetic Diversity	-	-	-	7.48 (0.374)	6.14 (0.456)
Mobility Index-Predicted Genetic Diversity Square	-	-	-	-6.12 (0.370)	-5.09 (0.447)
log Neolithic Transition Time	0.19 (0.112)	.22*** (0.000)	.22*** (0.000)	.27*** (0.000)	.26*** (0.000)
log Percentage Arable Land	.17** (0.014)	.05** (0.019)	.05** (0.018)	0.03 (0.287)	0.03 (0.287)
log Absolute Latitude	.08** (0.030)	.04* (0.085)	.04* (0.085)	.05* (0.053)	.05* (0.054)
log Land Suit. for Agriculture	-0.07 (0.156)	-.05** (0.032)	-.05** (0.031)	-0.03 (0.214)	-0.03 (0.220)
Kleibergen-Paap rk lm Statistic	-	41.2*** (0.000)	41.7*** (0.000)	30.278 (0.000)	29.469 (0.000)
Cragg-Donald Wald F Statistic	6.64^{δ}	$446.8^{\dagger\dagger\dagger}$ (11.04)	$353.7^{\dagger\dagger\dagger}$ (13.97)	151.996 (11.04)	145.476 (13.97)
Stock-Wright lm s Statistic	1.56 (0.458)	2.62 (0.623)	2.65 (0.753)	9.7** (0.046)	10.02* (0.075)
Hansen J-Statistic	exact. id.	1.97 (0.374)	2.08 (0.556)	7.81** (0.02)	8.477** (0.037)
Endog. Test	1.51 (0.219)	10.2*** (0.006)	16.46*** (0.000)	1.96 (0.375)	1.24 (0.537)
N	19	106	106	94	94

Note: Excluded instruments - (1) 2 instruments: first-stage fitted values of observed genetic diversity (square) and migratory distance from east Africa, equation is exactly identified; (2) and (4) 4 instruments: aerial distance from east Africa, aerial distance from east Africa (square), terrestrial distance from London and terrestrial distance from London (square); (3) and (5) 5 instruments: the same as in column (2) and (4) plus geodesic centroid latitude; Kleibergen-Paap rk lm statistic tests underidentification, under the null of the matrix of reduced form coefficients has rank=k1-1. Cragg-Donald Wald F statistic tests the null under which equation is weakly identified. this is compared with the Stock-Yogo weak id test critical values, which are reported in parentheses in that line. stock-wright lm s statistic tests the null under which the joint endogenous regressors have null coefficients. Hansen J-statistic tests the null under which the instruments are valid, i.e., uncorrelated with the error term.

Level of significance: *** for p-value<0.01; **for p-value<0.05;* for p-value<0.1. δ in column (1) 4.58 is the critical value for 15% critical iv size; † † † in columns (2) and (4), 11.04 is the critical value for relative IV bias of 5% (of the OLS bias); in columns (3) and (5), 13.97 is the critical value for relative IV bias of 5% (of the OLS bias).

values inside parentheses are p-values, except for the Cragg-Donald Wald F statistic. all regressions include continent dummies.

Table B.2.5: Technological Adoption and Genetic Diversity (2SLS Estimates)

	(1)	(2)	(3)	(4)	(5)
Dependent var.	*ind*	*ind*	*ind*	*ind*	*ind*
Observed Diversity	20.41 (0.553)	-	-	-	-
Observed Diversity Square	-15.96 (0.572)	-	-	-	-
Predicted Diversity	-	24.38** (0.010)	25.39*** (0.007)	-	-
Predicted Diversity Square	-	-17.26** (0.015)	-17.95** (0.011)	-	-
Mobility Index-Predicted Genetic Diversity	-	-	-	7.14* (0.083)	7.9** (0.048)
Mobility Index-Predicted Genetic Diversity Square	-	-	-	-5.71* (0.082)	-6.3* (0.050)
log Neolithic Transition Time	0.11 (0.463)	.21*** (0.000)	.21*** (0.000)	.13*** (0.001)	.13*** (0.001)
log Percentage Arable Land	.14* (0.072)	0.04 (0.263)	0.04 (0.267)	0.01 (0.606)	0.01 (0.606)
log Absolute Latitude	0.04 (0.355)	0.01 (0.610)	0.01 (0.607)	0.01 (0.545)	0.01 (0.555)
log Land Suit. for Agriculture	-.14** (0.011)	-.05* (0.070)	-.05* (0.070)	-.04** (0.038)	-.04** (0.039)
Kleibergen-Paap rk lm Statistic	-	41.2*** (0.000)	41.7*** (0.000)	30.278 (0.000)	29.469 (0.000)
Cragg-Donald Wald F Statistic	6.64^{δ}	$446.8^{\dagger\dagger\dagger}$ (11.04)	$353.7^{\dagger\dagger\dagger}$ (13.97)	151.996 (11.04)	145.476 (13.97)
Stock-Wright lm s Statistic	0.48 (0.787)	7.63 (0.106)	10.61* (0.06)	5.89 (0.207)	7.37 (0.194)
Hansen J-Statistic	exact. id.	5.14* (0.077)	6.10 (0.107)	2.55 (0.280)	2.71 (0.438)
Endog. Test	1.73 (0.188)	0.83 (0.659)	5.63* (0.06)	4.49 (0.106)	6.96** (0.031)
N	19	106	106	94	94

Note: Excluded instruments - (1) 2 instruments: first-stage fitted values of observed genetic diversity (square) and migratory distance from east Africa, equation is exactly identified; (2) and (4) 4 instruments: aerial distance from east Africa, aerial distance from east Africa (square), terrestrial distance from London and terrestrial distance from London (square); (3) and (5) 5 instruments: the same as in column (2) and (4) plus geodesic centroid latitude; Kleibergen-Paap rk lm statistic tests underidentification, under the null of the matrix of reduced form coefficients has rank=k1-1. Cragg-Donald Wald F statistic tests the null under which equation is weakly identified. this is compared with the Stock-Yogo weak id test critical values, which are reported in parentheses in that line. stock-wright lm s statistic tests the null under which the joint endogenous regressors have null coefficients. Hansen J-statistic tests the null under which the instruments are valid, i.e., uncorrelated with the error term.

Level of significance: *** for p-value<0.01; **for p-value<0.05;* for p-value<0.1. δ in column (1) 4.58 is the critical value for 15% critical iv size; † † † in columns (2) and (4), 11.04 is the critical value for relative IV bias of 5% (of the OLS bias); in columns (3) and (5), 13.97 is the critical value for relative IV bias of 5% (of the OLS bias).

values inside parentheses are p-values, except for the Cragg-Donald Wald F statistic. All regressions include continent dummies.

Appendix C

Appendix of Chapter 4

C.1 Technologies (Invention Dates) and Sources

Techs	Description	Year	Innovator	Source
ag_harvester	self-propelled machines that reap and thresh in one operation	1912	USA	Comin, D. & Mestieri, M. (2013). "If Technology Has Arrived Everywhere , Why Has Income Diverged?" INET Research Notes 26.
ag_milkingmachine	installations consisting of several complete milking units	1878	USA	Burton, L. D. V. (2010). Agriscience: Fundamentals and applications. Clifton Park, NY: Delmar Cengage Learning.
ag_tractor	wheel and crawler tractors (excluding garden tractors)	1892	USA	Comin, D. & Mestieri, M. (2013). "If Technology Has Arrived Everywhere , Why Has Income Diverged?" INET Research Notes 26.
atm	electromechanical devices that permit authorized users, typically using machine-readable plastic cards, to withdraw cash from their accounts and/or access other services	1960	USA	Simjian, L. (1963). Patent N.Âº 3079603. United States of America.
aviationpkm	Civil aviation passenger-KM traveled on scheduled services by companies registered in the country concerned. Not a measure of travel through a country's airports	1903	USA	Comin, D. & Mestieri, M. (2013). "If Technology Has Arrived Everywhere , Why Has Income Diverged?" INET Research Notes 26.
aviationtkm	Civil aviation ton-KM of cargo carried on scheduled services by companies registered in the country concerned. Not a measure of travel through a country's airports	1903	USA	Comin, D. & Mestieri, M. (2013). "If Technology Has Arrived Everywhere , Why Has Income Diverged?" INET Research Notes 26.
bed_acute	beds available for those seeking in-patient acute care, including diagnosis or treatment of an injury or illness and performance of surgery	1874	USA	National Association of Bedding Manufacturers, March 1964, Nation's Oldest Family-Held Bedding Firm: Adam Wuest, Inc.
bed_hosp	beds, including inpatient beds available in public, private, general, and specialized hospitals and rehabilitation centers. In most cases beds for both acute and chronic care are included	1874	USA	National Association of Bedding Manufacturers, March 1964, Nation's Oldest Family-Held Bedding Firm: Adam Wuest, Inc.
bed_longterm	beds for people who need assistance on a continuing basis due to chronic impairments and a reduced degree of independence in activities of daily living (including those in both hospitals and nursing homes)	1874	USA	National Association of Bedding Manufacturers, March 1964, Nation"s Oldest Family-Held Bedding Firm: Adam Wuest, Inc.
cabletv	Number of households that subscribe to a multi-channel television service delivered by a fixed line connection	1948	USA	Hitchner, J. R. (2010). Financial valuation: Applications and models. Hoboken, N.J: Wiley.
cellphone	Number of users of portable cell phones	1973	USA	Comin, D. & Mestieri, M. (2013). "If Technology Has Arrived Everywhere , Why Has Income Diverged?" INET Research Notes 26.

Continued on next page

Table C.1.1 – continued from previous page				
Techs	**Description**	**Year**	**Innovator**	**Source**
cheque	Number of payments by cheque (in millions)	1717	UK	Cheque & Credit Clearing Company. History of the Cheque: Cheque & Credit Clearing Company.[1]
computer	Number of self-contained computers designed for use by one person	1973	France	Comin, D. & Mestieri, M. (2013). "If Technology Has Arrived Everywhere , Why Has Income Diverged?" INET Research Notes 26.
creditdebit	Payments by credit and debit cards (in millions)	1949	USA	Bulliet, R. W. (1998). The Columbia history of the 20th century. New York: Columbia University Press.
eft	Number of transactions using payment cards at points of service (retail locations)	1949	USA	Bulliet, R. W. (1998). The Columbia history of the 20th century. New York: Columbia University Press.
elecprod	Gross output of electric energy (inclusive of electricity consumed in power stations) in KwHr	1882	USA	Comin, D. & Mestieri, M. (2013). "If Technology Has Arrived Everywhere , Why Has Income Diverged?" INET Research Notes 26.
fert_total	Metric tons of fertilizer consumed. Aggregate of 25 individual types listed in source	1910	Germany	Comin, D. & Mestieri, M. (2013). "If Technology Has Arrived Everywhere , Why Has Income Diverged?" INET Research Notes 26.
internetuser	access to the worldwide network	1983	USA	Comin, D. & Mestieri, M. (2013). "If Technology Has Arrived Everywhere , Why Has Income Diverged?" INET Research Notes 26.
kidney_dialpat	patients receiving dialysis treatments, both at centers and at home	1945	Netherlands	Ronco, C., Bellomo, R., & Kellum, J. A. (2009). Critical care nephrology. Philadelphia: Saunders/Elsevier.
kidney_homedialpat	patients receiving dialysis treatments at home	1962	Japan	Ing, T. S., Rahman, M. A., & Kjellstrand, C. M. (2012). Dialysis: History, development, and promise. Singapore: World Scientific.
loom_auto	operable looms (of a certain size) in place at year end and are either automatic or have automatic attachments (as opposed to ordinary looms)	1924	Japan	Mosk, C. (2007). Japanese Economic Development: Markets, Norms, Structures. New York: Routledge.
loom_total	operable looms in place at year end, including those that are automatic (as defined above) and those that are ordinary.	1924	Japan	Mosk, C. (2007). Japanese Economic Development: Markets, Norms, Structures. New York: Routledge.
mail	items mailed/received, with internal items counted one and cross-border items counted once for each country. May or may not include newspapers sent by mail, registered mail, or parcel post	1840	UK	Comin, D. & Mestieri, M. (2013). "If Technology Has Arrived Everywhere , Why Has Income Diverged?" INET Research Notes 26.
med_catscanner	computed tomography (CT) scanners, also known as "CAT" scans for computed axial tomography	1972	UK	Alshibli, K. & Reed, A. (2010) Advances in Computed Tomography for Geomaterials. London: ISTE Ltd.
med_lithotriptor	extracorporeal shock wave lithotripters, a machine typically used to break down kidney stones	1980	Germany	Nakada, S. Y., & Pearle, M. S. (2013). Surgical management of urolithiasis: Percutaneous, shockwave and ureteroscopy. New York, NY: Springer.
med_mammograph	dedicated mammography machines	1966	France	Karellasa, A. & Vedantham, S. (2008). Breast cancer imaging: A perspective for the next decade, Medical Physics, 35(11): 4878–4897.

Continued on next page

[1] Accessed on 1st of November of 2013 from: http://www.chequeandcredit.co.uk/cheque_and_credit_clearing/history_of_the_cheque/from_handwritten_to_printed_cheques/

Table C.1.1 – continued from previous page				
Techs	**Description**	**Year**	**Innovator**	**Source**
med_mriunit	magnetic resonance imaging (MRI) units	1977	USA	Placidi, D. (2012). MRI: Essentials for Innovative Technologies. New York: CRC Press.
med_radiationequip	pieces of equipment for treatment with x-rays or radionuclide	1895	USA	Beyzadeoglu, M., Ozyigit, G., & Ebruli, C. (2010). Basic radiation oncology. Heidelberg: Springer.
newspaper	newspaper copies circulated daily. Note that there is a tendency for news circulation to be under-reported, since data for weekly and biweekly publications are not included	1605	France	Spira, J. B. (2011). Overload!: How too much information is hazardous to your organization. Hoboken, N.J: Wiley.
pctdaysurg_cataract	cataract surgeries performed without a hospital stay	1967	USA	Yearly, P. (2005). They Were Giants 2005. Lincoln: iUniverse
pctdaysurg_cholecyst	cholecystectomies performed without a hospital stay	1882	Germany	Norton, J. A. (2008). Surgery: Basic science and clinical evidence. New York, NY: Springer.
pctdaysurg_hernia	hernia procedures performed without a hospital stay	1982	Germany	Schumpelick, V., & Fitzgibbons, R. J. (2007). Recurrent hernia: Prevention and treatment. Heidelberg: Springer Medizin.
pctdaysurg_lapcholecyst	laparoscopic cholecystectomies performed without a hospital stay	1985	Germany	Reynolds, W. (2001). The First Laparoscopic Cholecystectomy. JSLS, 5(1): 89–94.
pcthomedialysis	dialysis patients who receive treatment at home	1963	USA	Blagg, C. (2006). It's Time to Look at Home Hemodialysis in a New Light. Hemodialysis Horizons: Patient Safety & Approaches to Reducing Errors: 22–28.
pctimmunizdpt	children aged 12-23 months who received a DPT immunization (including all three doses) before the age of one year	1942	USA	Institute of Medicine (U.S.)., Howson, C. P., Howe, C. J., & Fineberg, H. V. (1991).Adverse effects of pertussis and rubella vaccines: A report of the Committee to Review the Adverse Consequences of Pertussis and Rubella Vaccines. Washington, D.C: National Academy Press.
pctimmunizmeas	children aged 12-23 months who received a measles immunization (one dose only) before the age of one year	1963	USA	Ndhlovu, Z. M. (2009). Cellular immune responses to measles virus-infection and vaccination.Â (Order No. 3356972, The Johns Hopkins University).Â ProQuest Dissertations and Theses,Â , 219.
pos	retail locations at which payment cards can be used Note: Per-capita data was converted to level data using WORLD BANK (2007) population data	1949	USA	Bulliet, R. W. (1998). The Columbia history of the 20th century. New York: Columbia University Press.
radio	Number of radios	1896	Russia	Ilcev, S. D. (2005). Global mobile satellite communication for maritime, land, and aeronautical applications. Dordrecht: Springer.
railline	Geographical/route lengths of line open at the end of the year. Narrow gauge lines generally included, but mountain railways, purely industrial lines not open to the public, and urban systems generally excluded	1825	UK	Comin, D. & Mestieri, M. (2013). "If Technology Has Arrived Everywhere , Why Has Income Diverged?" INET Research Notes 26.
railp	passenger journeys by railway. Free passengers typically excluded but may be included for some countries	1825	UK	Comin, D. & Mestieri, M. (2013). "If Technology Has Arrived Everywhere , Why Has Income Diverged?" INET Research Notes 26.
railpkm	Passenger journeys by railway in passenger-KM. Free passengers typically excluded but may be included for some countries	1825	UK	Comin, D. & Mestieri, M. (2013). "If Technology Has Arrived Everywhere , Why Has Income Diverged?" INET Research Notes 26.

Continued on next page

Table C.1.1 – continued from previous page				
Techs	**Description**	**Year**	**Innovator**	**Source**
railt	freight carried on railways (excluding livestock and passenger baggage). Freight for servicing of railroads is typically excluded but may be included for some countries	1825	UK	Comin, D. & Mestieri, M. (2013). "If Technology Has Arrived Everywhere , Why Has Income Diverged?" INET Research Notes 26.
railtkm	freight carried on railways (excluding livestock and passenger baggage). Freight for servicing of railroads is typically excluded but may be included for some countries	1825	UK	Comin, D. & Mestieri, M. (2013). "If Technology Has Arrived Everywhere , Why Has Income Diverged?" INET Research Notes 26.
ship_motor	motor ships (above a minimum weight) in use at midyear. Please see also general note on all ship-related series at end of list	1886	Germany	Guetat, G., & Ledru, E. (1997). Classic speedboats, 1916-1939. Osceola, WI: Motorbooks International.
ship_steam	steam ships (above a minimum weight) in use at midyear	1788	USA	Comin, D. & Mestieri, M. (2013). "If Technology Has Arrived Everywhere , Why Has Income Diverged?" INET Research Notes 26.
ship_steammotor	steam and motor ships (above a minimum weight) in use at midyear	1788	USA	Inventors. The history of steamboats.[2]
shipton_motor	motor ships (above a minimum weight) in use at midyear	1886	Germany	Comin, D. & Mestieri, M. (2013). "If Technology Has Arrived Everywhere , Why Has Income Diverged?" INET Research Notes 26.
shipton_steam	steam ships (above a minimum weight) in use at midyear	1788	USA	Inventors. The history of steamboats.[2]
shipton_steammotor	steam and motor ships (above a minimum weight) in use at midyear	1788	USA	Inventors. The history of steamboats.[2]
spindle_mule	mule spindles in place at year end	1779	UK	Comin, D. & Mestieri, M. (2013). "If Technology Has Arrived Everywhere , Why Has Income Diverged?" INET Research Notes 26.
spindle_ring	ring spindles in place at year end	1828	USA	Wallace, A. F. C. (2005). Rockdale: The growth of an American village in the early Industrial Revolution. Lincoln: University of Nebraska Press.
steel_acidbess	Crude steel production (in metric tons) by the acid Bessemer process (an early steel process)	1855	UK	Gasik, M. (2013). Handbook of Ferroalloys: Theory and Technology. Butterworth-Heinemann.
steel_basicbess	Crude steel production (in metric tons) by the basic Bessemer process (an early steel process)	1878	UK	Almqvist, E. (2003). History of industrial gases. New York, N.Y: Kluwer Academic/Plenum Publishers.
steel_bof	Crude steel production (in metric tons) in blast oxygen furnaces (a process that replaced Bessemer and OHF processes)	1952	Austria	Comin, D. & Mestieri, M. (2013). "If Technology Has Arrived Everywhere , Why Has Income Diverged?" INET Research Notes 26.
steel_eaf	Crude steel production (in metric tons) in electric arc furnaces (a process that complemented and improved upon Bessemer and OHF processes)	1907	USA	Comin, D. & Mestieri, M. (2013). "If Technology Has Arrived Everywhere , Why Has Income Diverged?" INET Research Notes 26.
steel_ohf	Crude steel production (in metric tons) in open hearth furnaces (a process that complemented the Bessemer process)	1865	UK	Comin, D. & Mestieri, M. (2013). "If Technology Has Arrived Everywhere , Why Has Income Diverged?" INET Research Notes 26.
steel_other	Crude steel production (in metric tons) by methods other than those listed here	1614	UK	McCosh, F. W. J. (1984). Boussingault, chemist and agriculturist. Dordrecht: D. Reidel Pub. Co.

Continued on next page

[2] Accessed on 1st of November of 2013 from: http://inventors.about.com/library/inventors/blsteamship.htm

Table C.1.1 – continued from previous page				
Techs	**Description**	**Year**	**Innovator**	**Source**
steel_stainless	Stainless steel production (in metric tons). Stainless and crude steel have different functions	1904	France	Reardon, A. C. (2011). Metallurgy for the non-metallurgist. Materials Park, Ohio: ASM International.
surg_appendectomy	Number of appendectomies performed	1735	UK	Stockman, J. (2013). Year Book of Pediatrics 2013: Pediatrics. London: Elsevier Health Sciences.
surg_breastcnsv	breast conservation surgeries performed	1976	USA	Ueno, N. T., & Cristofanilli, M. (2012). Inflammatory breast cancer: An update. Dordrecht: Springer.
surg_cardcath	cardiac catheterizations (insertion of a catheter into a chamber or vessel of the heart) performed	1929	Germany	Lilly, L. S., & Harvard Medical School. (2011). Pathophysiology of heart disease: A collaborative project of medical students and faculty. Baltimore, MD: Wolters Kluwer/Lippincott Williams & Wilkins.
surg_cholecyst	cholecystectomies (gallbladder removals) performed, either laparoscopically or by other methods	1882	Germany	Norton, J. A. (2008). Surgery: Basic science and clinical evidence. New York, NY: Springer.
surg_corbypass	coronary bypass surgeries performed	1960	USA	DeSilva, R. (2013). Heart disease. Santa Barbara, Calif: Greenwood.
surg_corinterven	percutaneous coronary interventions (used to reduced or eliminate the symptoms of coronary artery disease) performed	1977	Switzerland	Estafanous, F. G., Barash, P. G., & Reves, J. G. (2001). Cardiac anesthesia: Principles and clinical practice. Philadelphia: Lippincott Williams & Wilkins.
surg_corstent	coronary stenting procedures performed. This is a particular type of percutaneous coronary intervention	1994	USA	In Bandhyopadhya, A., & In Bose, S. (2013). Characterization of biomaterials.
surg_hernia	procedures performed to correct inguinal and femoral hernias (the most common types)	1879	UK	Hupp, F. (1924). Intra-abdominal rupture of intestine following strangulated femoral hernia. Ann. Surg. 80 (4): 504-10
surg_hipreplace	hip replacement surgeries performed	1891	Germany	Gomez, P. & Morcuende J. (2005). Early Attempts at Hip Arthroplasty. Iowa Orthop J. 25: 25-29.
surg_hysterectomy	vaginal hysterectomies performed (does not include abdominal or laparoscopic procedures)	1813	Germany	Mettler, L. (2007). Manual of new hysterectomy techniques. New Delhi: Jaypee Brothers Med. Publ. [u.a..
surg_kneereplace	knee replacement surgeries	1968	UK	Scuderi, G. R., & Tria, A. J. (2002). Surgical techniques in total knee arthroplasty. New York: Springer.
surg_lapcholecyst	cholecystectomies (gallbladder removals) performed laparoscopically	1882	Germany	Norton, J. A. (2008). Surgery: Basic science and clinical evidence. New York, NY: Springer.
surg_mastectomy	mastectomies performed	1882	USA	Pilnik, S. (2003). Common breast lesions: A photographic guide to diagnosis and treatment. Cambridge: Cambridge University Press.
surg_pacemaker	pacemaker implantation procedures performed	1926	Australia	Torok, S. & Holper, P. (2006). Inventing Millions: Creating wealth, changing lives. New Delhi: Orient Paperbacks
telegram	telegrams sent	1835	USA	Comin, D. & Mestieri, M. (2013). "If Technology Has Arrived Everywhere , Why Has Income Diverged?" INET Research Notes 26.
telephone	mainline telephone lines connecting a customer's equipment to the public switched telephone network as of year end	1876	USA	Comin, D. & Mestieri, M. (2013). "If Technology Has Arrived Everywhere , Why Has Income Diverged?" INET Research Notes 26.

Continued on next page

Table C.1.1 – continued from previous page				
Techs	**Description**	**Year**	**Innovator**	**Source**
transplant_bonemarrow	bone marrow transplants performed	1956	USA	Kidder, D. S., Oppenheim, N. D., & Young, B. K. (2009). The intellectual devotional health: Revive your mind, complete your education, and digest a daily dose of wellness wisdom. Emmaus, Pa.: Rodale.
transplant_heart	heart transplants performed	1968	USA	Comin, D. & Mestieri, M. (2013). "If Technology Has Arrived Everywhere , Why Has Income Diverged?" INET Research Notes 26.
transplant_kidney	kidney transplants performed	1954	USA	Comin, D. & Mestieri, M. (2013). "If Technology Has Arrived Everywhere , Why Has Income Diverged?" INET Research Notes 26.
transplant_liver	liver transplants performed	1963	USA	Comin, D. & Mestieri, M. (2013). "If Technology Has Arrived Everywhere , Why Has Income Diverged?" INET Research Notes 26.
transplant_lung	lung transplants performed.	1963	USA	Couture, K. A., & Couture, K. A. (2001). The lung transplantation handbook. Victoria, B.C: Trafford.
tv	television sets in use	1884	Germany	Peddie, J. (2013). The history of visual magic in computers: How beautiful images are made in CAD, 3D, VR and AR. London: Springer.
txtlmat_artif	artificial (cellulosic) fibers used in spindles	1865	UK	Baird, G., Mertins, D., & Mies, . R. L. (1994). The presence of Mies. New York, NY: Princeton Architectural Press.
txtlmat_synth	synthetic (non-cellulosic) fibers used in spindles	1924	USA	Comin, D. & Mestieri, M. (2013). "If Technology Has Arrived Everywhere , Why Has Income Diverged?" INET Research Notes 26.
vehicle_car	passenger cars (excluding tractors and similar vehicles) in use. Numbers typically derived from registration and licensing records, meaning that vehicles out of use may occasionally be included.	1885	Germany	Comin, D. & Mestieri, M. (2013). "If Technology Has Arrived Everywhere , Why Has Income Diverged?" INET Research Notes 26.
vehicle_com	commercial vehicles, typically including buses and taxis (excluding tractors and similar vehicles), in use. Numbers typically derived from) registration and licensing records, meaning that vehicles out of use may occasionally be included	1885	Germany	Comin, D. & Mestieri, M. (2013). "If Technology Has Arrived Everywhere , Why Has Income Diverged?" INET Research Notes 26.

Appendix D

Appendix of Chapter 5

D.1 Appendix: Additional Unit Root and Cointegration Tests

Table D.1.1: Panel Unit-Root tests (differenced variables)

		(1)	(2)	(3)	(4)	(5)
Variable	Lag	Gini Net Income SWIID (>30)	Gini Net Income SWIID (>30, ./sd)	Human Capital	TFP	Openness
				Pesaran (2007) Test without Trend		
Zt-stat	0	-19.818***	-32.761***	-0.893	-41.383***	-54.611***
p-value		(0.000)	(0.000)	(0.186)	(0.000)	(0.000)
Zt-stat	1	-9.705***	-26.800***	-2.299**	-26.245***	-45.136***
p-value		(0.000)	(0.000)	(0.011)	(0.000)	(0.000)
Zt-stat	2	-12.056***	-16.646***	-3.550***	-18.507***	-30.995***
p-value		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Zt-stat	3	-8.450***	-12.134***	-5.837***	-13.252***	-23.578***
p-value		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
				Pesaran (2007) Test with Trend		
Zt-stat	0	-18.382***	-30.603***	2.477	-40.184***	-53.027***
p-value		(0.000)	(0.000)	(0.993)	(0.000)	(0.000)
Zt-stat	1	-6.394***	-23.520***	1.140	-23.809***	-41.989***
p-value		(0.000)	(0.000)	(0.873)	(0.000)	(0.000)
Zt-stat	2	-9.161***	-12.366***	-0.112	-15.671***	-26.970***
p-value		(0.000)	(0.000)	(0.455)	(0.000)	(0.000)
Zt-stat	3	-7.046***	-8.002***	-2.620***	-10.564	-19.422***
p-value		(0.000)	(0.000)	(0.004)	(0.000)	(0.000)
Number of Countries		82	82	128	106	155
N. of Observations		2992	2992	6566	4888	7605
Avr. N. of Obs.		37.8	37.8	54.4	50.6	52.9

Note: All variables are in natural logarithms. >30 indicates that only cross-sections with more that 30 time-series observations are included. ./sd indicates when the Gini coefficient is divided by the source standard-deviation to account for data uncertainty. Level of significance: *** for p-value<0.01; **for p-value<0.05;* for p-value<0.1.

Table D.1.2: Cointegration tests

	Lag	(1) Trend	(5) Test Gt	(6) Test Ga	(7) Test Pt	(8) Test Pa
Dependent Variable			Gini Coefficient net income (>30) (from SIIWD)			
	1	No	-2.649***	-6.297	-13.782***	-8.885***
p-value			(0.000)	(0.767)	(0.000)	(0.000)
	1	Yes	-3.567***	-8.984	-15.144***	-13.135***
p-value			(0.000)	(0.982)	(0.000)	(0.001)
	2	No	-2.745***	-6.394	-10.735***	-5.930**
p-value			(0.000)	(0.740)	(0.000)	(0.036)
	2	Yes	-3.201***	-8.194	-12.938***	-8.745
p-value			(0.000)	(0.996)	(0.000)	(0.557)
Dependent Variable			Human Capital (from PWT 8.0)			
	1	No	-2.037*	-4.068	-8.556**	-2.302
p-value			(0.088)	(0.996)	(0.038)	(0.979)
	1	Yes	-1.964	-8.196	-9.005	-6.354
p-value			(0.990)	(0.996)	(0.849)	(0.976)
	2	No	-2.096**	-3.763	-6.934	-1.961
p-value			(0.048)	(0.998)	(0.442)	(0.992)
	2	Yes	-1.797	-7.665	-8.186	-6.024
p-value			(1.000)	(0.999)	(0.976)	(0.987)

Note: All variables are in natural logarithms. >30 indicates that only cross-sections with more that 30 time-series observations are included. All tests include a constant. Level of significance: *** for p-value<0.01; **for p-value<0.05;* for p-value<0.1. Rejection of H0 in Ga and Gt tests should be taken as evidence of cointegration of at least one of the cross-sectional units. Rejection of H0 in Pa and Pt tests should therefore be taken as evidence of cointegration for the panel as a whole.

D.2 Appendix: Lists of Countries

This section lists the countries used in the main regressions in the paper (Tables 5.5 - columns (3) and (4), Table 5.7).

D.2.1 Sample in Tables 5.5, column (3)

Argentina, Armenia, Australia, Austria, Barbados, Belgium, Bolivia, Botswana, Brazil, Bulgaria, Burundi, Cameroon, Canada, Central African Republic, Chile, China, Colombia, Costa Rica, Cote d'Ivoire, Croatia, Cyprus, Czech Republic, Denmark, Dominican Republic, Ecuador, Egypt, Estonia, Fiji, Finland, France, Germany, Greece, Guatemala, Honduras, Hong Kong, Hungary, Iceland, India, Indonesia, Iran, Ireland, Israel, Italy, Jamaica, Japan, Jordan, Kazakhstan, Kenya, Republic of Korea, Kyrgyz Republic, Latvia, Lesotho, Lithuania, Luxembourg, Malaysia, Malta, Mauritania, Mauritius, Mexico, Moldova, Mongolia, Morocco, Mozambique, Namibia, Netherlands, New Zealand, Niger, Norway, Panama, Paraguay, Peru, Philippines, Poland, Portugal, Romania, Russian Federation, Rwanda, Senegal, Serbia, Sierra Leone, Singapore, Slovak Republic, Slovenia, South Africa, Spain, Sri Lanka, Swaziland, Sweden, Switzerland, Taiwan, Tajikistan, Tanzania, Thailand, Togo, Trinidad and Tobago, Tunisia, Turkey, Ukraine, United Kingdom, United States, Uruguay, Venezuela, Zimbabwe.

D.2.2 Sample in Tables 5, column (4), and Table 8, column (1)

Argentina, Australia, Belgium, Brazil, Bulgaria, Canada, Chile, China, Colombia, Costa Rica, Cote d'Ivoire, Denmark, Egypt, Estonia, Fiji, Finland, France, Germany, Greece, Hong Kong, Hungary, India, Indonesia, Iran, Ireland, Israel, Italy, Jamaica, Japan, Jordan, Kazakhstan, Kenya, Republic of Korea, Kyrgyz Republic, Latvia, Lithuania, Malaysia, Mauritius, Mexico, Moldova, Morocco, Netherlands, New Zealand, Norway, Panama, Peru, Philippines, Poland, Portugal, Russian Federation, Sierra Leone, Singapore, South Africa, Spain, Sri Lanka, Sweden, Switzerland, Taiwan, Tanzania, Thailand, Trinidad and Tobago, Tunisia, Turkey, Ukraine, United Kingdom, United States, Uruguay, Venezuela.

D.2.3 Sample in Tables 5.7, columns (2), (3), (4) and (5)

Argentina, Australia, Bangladesh, Belgium, Brazil, Bulgaria, Canada, Chile, China, Colombia, Costa Rica, Cote d'Ivoire, Denmark, Egypt, El Salvador, Estonia, Fiji, Finland, France, Germany, Greece, Hong Kong, Hungary, India, Indonesia, Iran, Ireland, Israel, Italy, Jamaica, Japan, Jordan, Kazakhstan, Kenya, Republic of Korea, Kyrgyz Republic, Latvia, Lithuania, Malawi, Malaysia, Mauritius, Mexico, Moldova, Morocco, Nepal, Netherlands, New Zealand, Norway, Pakistan, Panama, Peru, Philippines, Poland, Portugal, Russian Federation, Sierra Leone, Singapore, South Africa, Spain, Sri Lanka, Sweden, Switzerland, Taiwan, Tanzania, Thailand, Trinidad and Tobago, Tunisia, Turkey, Ukraine, United Kingdom, United States, Uruguay, Venezuela, Zambia.

D.2.4 Sample in Tables 5.7, column (6)

Argentina, Australia, Bangladesh, Belgium, Brazil, Bulgaria, Canada, Chile, China, Colombia, Costa Rica, Cote d'Ivoire, Denmark, Egypt, El Salvador, Fiji, Finland, France, Germany, Greece, Hong Kong, Hungary, India, Indonesia, Iran, Ireland, Israel, Italy, Jamaica, Japan, Jordan, Kenya, Republic of Korea, Malawi, Malaysia, Mauritius, Mexico, Morocco, Nepal, Netherlands, New Zealand, Norway, Pakistan, Panama, Peru, Philippines, Poland, Portugal, Sierra Leone, Singapore, South Africa, Spain, Sri Lanka, Sweden, Switzerland, Taiwan, Tanzania, Thailand, Trinidad and Tobago, Tunisia, Turkey, United Kingdom, United States, Uruguay, Venezuela, Zambia.

D.3 Appendix: Alternative Corrected Gini index

Table D.3.1: Inequality, Human Capital, TFP, and Openness

	(1)	(2)
Dependent Variable: Gini Measure	Gini Net post-tax; post-transfer ./(1+sd)	Gini Net post-tax; post-transfer ./(1+sd),>30
$hcap$	1.10***	1.44***
	(.004)	(.000)
TFP	.006	-0.058
	(.931)	(0.498)
$Open$	.02	0.02
	(.460)	(0.461)
N Observ.	3300	2593
Avr. N Obs.	32	38.1
Min-Max	7-52	21-52
Number Countries	103	68
Wald	9.01**	15.39***
CD-test (res)	-	1.10 (0.272)
Stat-test (res)	-	rejects I(1)
sig. signs /countries for $hcap$	↗(38)↘(12)	↗(31)↘(4)
sig. signs /countries for TFP	↗(19)↘(19)	↗(11)↘(16)
sig. signs /countries for $Open$	↗(17)↘(5)	↗(15)↘(4)

Note: Dependent Variables natural logarithm of the Gini coefficient. All variables are in natural logarithms. A constant is included in the regressions but omitted from the Table. Values between parentheses below coefficients are p-values from robust (clustered) standard errors. Level of significance: *** for p-value<0.01; **for p-value<0.05;* for p-value<0.1. Wald test is a joint significance test for the regressors. CD-test is a Pesaran (2004) cross-section dependence test on the null of cross-section independence done on the residuals from the regression (p-value presented between parentheses). Stat-test is the Pesaran (2007) unit root test made on the residuals. This test used 3 lags and rejects I(1) means that in all lags the test of unit root rejects. The lists of countries that enter in columns (3) and (4) are provided in the Appendix B. ./(1+sd) indicates when the Gini coefficient is divided by 1 plus the source standard-deviation to account for data uncertainty.

Table D.3.2: Inequality, Human Capital, TFP, and Openness (Rich versus Poor countries)

	Rich Sample		Poor Sample	
	(1)	(2)	(3)	(4)
Dependent Variable: Gini Measure	Gini Net post-tax; post-transfer ./(1+sd)	Gini Net post-tax; post-transfer (>30, ./(1+sd))	Gini Net post-tax; post-transfer ./(1+sd)	Gini Net post-tax; post-transfer (>30, ./(1+sd))
$hcap$	1.49***	1.29***	0.76	0.499
	(0.003)	(0.004)	(0.170)	(0.451)
TFP	0.02	-0.03	-0.046	-0.097
	(0.853)	(0.792)	(0.563)	(0.396)
$Open$	0.01	-0.04	0.024	-0.016
	(0.777)	(0.534)	(0.395)	(0.712)
N Observ.	1657	1431	1643	1162
Avr. N Obs.	36.8	40.9	28.3	35.2
Min-Max	12-52	22-52	7-48	21-48
Number Countries	45	35	58	33
Wald	9.22**	8.62**	2.94	1.43
CD-test (res)	-	1.02 (0.307)	-	-0.06 (0.955)
Stat-test (res)	-	reject I(1)	-	reject I(1)

Note: Note: Dependent Variables natural logarithm of the Gini coefficient. All variables are in natural logarithms. A constant is included in the regressions but omitted from the Table. Values between parentheses below coefficients are p-values from robust (clustered) standard errors. Level of significance: *** for p-value<0.01; **for p-value<0.05;* for p-value<0.1. Wald test is a joint significance test for the regressors. CD-test is a Pesaran (2004) cross-section dependence test on the null of cross-section independence done on the residuals from the regression (p-value presented between parentheses). Stat-test is the Pesaran (2007) unit root test made on the residuals. This test used 3 lags and rejects I(1) means that in all lags the test of unit root rejects. The lists of countries that enter in columns (3) and (4) are provided in the Appendix B. ./(1+sd) indicates when the Gini coefficient is divided by 1 plus the source standard-deviation to account for data uncertainty.

Table D.3.3: Inequality, Human Capital, TFP, and Openness (Robustness)

Dependent Variable	Gini Coefficient net income (./(1+sd), >30)					
Vars. only as CS Avr.	Open	Open; TFP	Open; TFP; GDP p.c.	Open; without TFP	Open; TFP	Open; TFP
Lags of CS Avr.	0	0	0	0	2 (Gini); 3 (hcap); 0 (other)	3 (all)
	(1)	(2)	(3)	(4)	(5)	(6)
$hcap$	1.31*** (0.001)	1.54*** (0.000)	0.62** (0.035)	1.52*** (0.000)	1.10** (0.028)	2.23*** (0.009)
TFP	-0.064 (0.441)	-	-	-	-	-
N Observ.	2593	2855	2855	2855	2463	2445
Avr. N Obs. Min-Max	38.1 21-52	38.6 21-52	38.6 21-52	38.6 21-52	33.3 18-49	33.5 19-49
Number Countries	68	74	74	74	74	73
Wald	55.98***	68.92***	34.68***	34.68***	43.10***	24.65*
CD-test (res)	1.15 (0.250)	1.14 (0.254)	0.21 (0.834)	0.39 (0.694)	3.72*** (0.000)	6.41*** (0.000)
Stat-test (res)	Reject I(1)	Reject I(1)	Reject I(1)	Reject I(1)	Reject I(1)	Reject I(1)
sig. signs /countries for $hcap$	↗(13)↘(3)	↗(41)↘(9)	↗(19)↘(8)	↗(42)↘(9)	↗(14)↘(9)	↗(16)↘(11)
sig. signs /countries for TFP	↗(13)↘(18)	-	-	-	-	-

Note: Dependent Variables natural logarithm of the Gini coefficient. All variables are in natural logarithms. A constant is included in the regressions but omitted from the Table. Values between parentheses below coefficients are p-values from robust (clustered) standard errors. Level of significance: *** for p-value<0.01; **for p-value<0.05;* for p-value<0.1. Wald test is a joint significance test for the regressors. CD-test is a Pesaran (2004) cross-section dependence test on the null of cross-section independence done on the residuals from the regression (p-value presented between parentheses). Stat-test is the Pesaran (2007) unit root test made on the residuals. This test used 3 lags and rejects I(1) means that in all lags the test of unit root rejects. The lists of countries that enter in columns (3) and (4) are provided in the Appendix B. Vars. only as CS Avr. means variables that only enter regression as cross-section average but not as country-specific variable. ./(1+sd) indicates when the Gini coefficient is divided by 1 plus the source standard-deviation to account for data uncertainty.

Appendix E

Appendix of Chapter 6

E.1 List of Technologies and Abbreviations

- wheel and crawler tractors (excluding garden tractors); definition in the source: tractor; abbreviation: tt
- electromechanical devices that permit authorized users, typically using machine readable plastic cards, to withdraw cash from their accounts and/or access other services; definition in the source: atm; abbreviation: atm
- Civil aviation passenger-KM traveled on scheduled services by companies registered in the country concerned. Not a measure of travel through a country airports; definition in the source: aviationpkm; abbreviation: a1
- Civil aviation ton-KM of cargo carried on scheduled services by companies registered in the country concerned. Not a measure of travel through a country's airports; definition in the source: aviationtkm; abbreviation: a2
- Number of users of portable cell phones; definition in the source: cell phone; abbreviation: cp
- Number of self-contained computers designed for use by one person; definition in the source: computer; abbreviation: ct
- Gross output of electric energy (inclusive of electricity consumed in power stations) in KwHr; definition in the source: elecprod; abbreviation: ep
- access to the worldwide network; definition in the source: internetuser; abbreviation: it
- Number of Radios; definition in the source: radio; abbreviation: ra
- Geographical/route lengths of line open at the end of the year. Narrow gauge lines generally included, but mountain railways, purely industrial lines not open to the public, and urban systems generally excluded; definition in the source: railline; abbreviation: r1
- Passenger journeys by railway in passenger-KM. Free passengers typically excluded but may be included for some countries; definition in the source: railpkm; abbreviation: r2
- freight carried on railways (excluding livestock and passenger baggage). Freight for servicing of railroads is typically excluded but may be included for some countries; definition in the source: railtkm; abbreviation: r3
- steamships (above a minimum weight) in use at midyear; definition in the source: ship-ton-steammotor; abbreviation: sh
- Crude steel production (in metric tons) in blast oxygen furnaces (a process that replaced Bessemer and OHF processes); definition in the source: steel-bof; abbreviation: s1

- Crude steel production (in metric tons) in electric arc furnaces (a process that complemented and improved upon Bessemer and OHF processes); definition in the source: steel-eaf; abbreviation: s2
- Telegrams; definition in the source: telegram; abbreviation: tg
- mainline telephone lines connecting a customer's equipment to the public switched telephone network as of year end; definition in the source: telephone; abreviation: tl
- television sets in use; definition in the source: tv; abbreviation: tv
- passenger cars (excluding tractors and similar vehicles) in use. Numbers typically derived from registration and licensing records, meaning that vehicles out of use may occasionally be included.; definition in the source: vehicle-car; abbreviation: cr
- commercial vehicles, typically including buses and taxis (excluding tractors and similar vehicles), in use. Numbers typically derived from registration and licensing records, meaning that vehicles out of use may occasionally be included; definition in the source: vehicle-com; abbreviation: tr

E.2 Restricted Regressions

Table E.2.1: Alternative Restricted Regressions

	(1)	(2)	(3)	(4)
Estimator	Pesaran (2006)	Pesaran (2006)	Pesaran (2006)	Pesaran (2006)
Modern ICT	0.07***	-	-	-
	(0.001)	-	-	-
Older ICT	-	0.14***	-	-
	-	(0.000)	-	-
Production	-	-	0.07***	-
	-	-	(0.000)	-
Transportation	-	-	-	0.06***
	-	-	-	(0.000)
Wald	10.54***	25.12***	12.91***	18.3***
Avr. Obs.	9.6	17.6	23.4	17.4
N. Countries	33	50	43	33
Total Obs.	316	882	1006	574

Note: Dependent Variables natural logarithm of the Gini coefficient. A constant is included in regressions but omitted from the table. Values in parentheses are p-values. P-values of coefficients are based on robust standard errors. Level of significance: *** for p-value<0.01; **for p-value<0.05;* for p-value<0.1. Additional CS Avg means the additional variables added as controls as cross-section averages.