TESIS DE LA UNIVERSIDAD DE ZARAGOZA

2020

37

Cristina Martínez Asenjo

# Three essays on Convergence and Health Economics

Departamento
Análisis Económico

Director/es
Montañes Bernal, Antonio

http://zaguan.unizar.es/collection/Tesis

ISSN 2254-7606

Prensas de la Universidad
Universidad Zaragoza

ISSN 2254-7606

Tesis Doctoral

# THREE ESSAYS ON CONVERGENCE AND HEALTH ECONOMICS


Autor

Cristina Martínez Asenjo

Director/es

Montañes Bernal, Antonio

**UNIVERSIDAD DE ZARAGOZA**

Análisis Económico



2019

# Three essays on Convergence and Health Economics

A thesis submitted by

Cristina Martinez Asenjo

Thesis advisor

Antonio Montañés Bernal

For the degree of

Doctor of Philosophy in Economics

At Department of Economic Analysis

Universidad Zaragoza

2019

# *Acknowledgements*

Writing a dissertation is always a hard work but it become much more bearable when you have wonderful people around you. For this reason, although academically this is not the most relevant part of the Thesis, if it is personally.

Firstly, I want to acknowledge the continuous support of two exceedingly great people without whom this work really would not have been possible;

Antonio Montañés Bernal, my thesis supervisor, for his invaluable help, for having always encouragement words for me, for never giving up during the stoppages or when the deadlines lengthen and the circumstances become complex. And of course, for teaching me to enjoy econometrics!

Gabriel Alvarez Casanovas, my internist doctor, for always being the helping hand when the disease squeezes, for never telling me that I cannot, even though the reality is adverse.

Thanks you both for your human and professional integrity, for believing in me and help me fulfil this dream. The reward for the effort made has been to meet you both on the way.

My gratitude to all administrative staff of Zaragoza University, especially to the Analysis Economic Department; Angel y Maria Angeles for their availability and willingness, always with a smile for me. To Fernando Palacios for his high professional competence and his incalculable assistance in all the bureaucratic procedures in which we have been involved.

Special Thanks to Maria Navarro and Pierre Perron for their warm welcome and for giving me the opportunity to learn from you both during my visiting researches at Lancaster University and Boston University respectively.

In a very special way, I would like to thank my family and friends, my greatest treasure in the world, their constant support;

To Jose Carlos, my vital support, for his patience, his time and his dedication. For being the column when everything else collapses and rebuild the world with me every day. For his unconditional love without guarantees.  Thank you

To my parents for their always tireless fight, for believing I could achieve everything I want even when I have my doubts, and being able to see who I am.

To my friends; all those who have always been there for me, friends of old times and friends of now, to the new ones that I have been finding and also to those who have disappeared during the long journey of this thesis, because I learned something from everyone.

To Jayden, my little miracle, for showing me that everything is possible and giving me the strength to fight every day.

Finally, I cannot forget to thanks the appreciated financial support awarded by the Spanish Ministry of Education, Culture and Sport with FPU program (AP2010-3996), and the Science, Technology and University Department of DGA for the grant of research personnel training (B072/11).

To all of them my deepest gratitude for being part of this great project.

# *Agradecimientos*

Escribir una tesis siempre es un trabajo duro pero se hace mucho más llevadero cuando tienes gente maravillosa alrededor. Por ello, aunque académicamente esta no sea la más parte más relevante de la Tesis, si lo es a nivel personal.

En primer lugar quiero a gradecer el apoyo continuo de dos grandísimas personas sin las que realmente este trabajo no habría sido posible;

Antonio Montañés Bernal, mi director, por su inestimable ayuda y sus palabras de ánimo siempre, por no darse por vencido en los parones, cuando los plazos se alargan y las circunstancias se tornan complejas. Y por supuesto, por enseñarme a disfrutar con la econometría.

A Gabriel Álvarez Casanovas, mi médico internista, por ser siempre la mano amiga cuando la enfermedad aprieta, por no decirme nunca que no puedo, aunque la realidad se muestre adversa.

Gracias a ambos por vuestra integridad humana y profesional, por creer en mí y ayudarme a cumplir este sueño. La recompensa al esfuerzo realizado ha sido conoceros a vosotros en el camino.

Mi agradecimiento a todo el personal administrativo de la Universidad de Zaragoza, y especialmente del departamento de Análisis Económico; A Ángel y María Ángeles, siempre disponibles con una sonrisa. A Fernando Palacios, por su gran competencia profesional y su ayuda incalculable en todos los trámites administrativos en los que no hemos visto envueltos.

A María Navarro Paniagua y Pierre Perron por su cariñosa acogida y darme la oportunidad de aprender de ellos durante mis estancias en la Universidad de Lancaster y la Universidad de Boston respectivamente.

Agradecer de una manera muy especial a mi familia y amigos, mi mayor tesoro en el mundo, su apoyo constante;

A José Carlos, mi soporte vital, por su paciencia, su tiempo y su dedicación. Por ser la columna cuando todo lo demás se derrumba y reconstruir conmigo el mundo cada día, por su amor incondicional y sin garantías. Gracias

A mis padres por su lucha incansable siempre, por creer en mí incluso cuando yo he tenido mis propias dudas, y ser capaces siempre de ver quién soy.

A mis amigos, todos aquellos que siempre han estado ahí, a los de antes y a los de ahora, a los nuevos que he ido encontrando y también a los que han desaparecido durante el largo camino de esta tesis, porque de todos he aprendido algo.

A Jayden, mi pequeño milagro, por demostrarme que todo es posible y por darme las fuerzas para luchar cada día.

Por ultimo quiero agradecer el importante apoyo económico recibido por parte del Ministerio de Educación, cultura y deporte por la beca concedida bajo el programa FPU (AP2010-3996), y al Departamento de Ciencia, Tecnología y Universidad de la Diputación General de Aragón por la subvención otorgada para la formación de personal investigador (B072/11).

A todos ellos mi más sincero agradecimiento por formar parte de este gran proyecto.

# Contents

# List of Tables

Chapter 1:



Chapter 2:



Chapter 3:

# List of Figures

Chapter 1:



Chapter 2:



Chapter 3:

*"Health is a driver, an indicator and a result of sustainable development. Healthy people are better able to contribute to the development of their countries".*

United Nations Development Program (UNDP)

# Introduction

Health is one of the mainstays for the economic development and the welfare state of any society. Therefore, it comes as no surprise that the main international organizations, such as United Nations, World Bank or International Monetary Fund, have highlighted its importance by granting numerous projects worldwide and by creating international institutions dedicated exclusively to the field of health. This is the case of the World Health Organization (WHO), which was constituted in 1948. We can additionally emphasize the importance of health by taking into account the United Nations Millennium Declaration, the most ambitious project of this organization so far, signed in September 2000 by 189 member countries, in which they undertook to pursue eight Millennium Development Goals (MDGs) of the United Nations. These MDGs were replaced in 2016 by the Sustainable Development Goals (SDGs).

Although all the objectives set out in the MDGs have a direct or indirect relationship with health, there are three of them dedicated exclusively to it: MDG4- Reduce child mortality, MDG5- Improve maternal health and MDG6- Combat Human Immunodeficiency Virus (HIV) / Acquire Immunodeficiency Syndrome (AIDS), malaria and other diseases. Currently, these objectives have been included in objective 3 of the SDGs, denominated "Health and well-being", whose targets for 2030 include: To end the preventable deaths of newborns and children under 5 years of age, making all countries try to reduce neonatal mortality to at least as low as 12 per 1,000 live births and the mortality of children under 5 years old at least as low as 25 per 1,000 live births. End to the epidemics of AIDS, tuberculosis, malaria and neglected tropical diseases. As well as fighting hepatitis, waterborne diseases and other communicable diseases.

The interest that the field of health arouse in the economy has attracted a large number of researchers. Following the same interest, it seems relevant to analyze how and to what degree have the objectives established by the United Nations in the MDGs been achieved. Specifically, we analyze the evolution of the HIV / AIDS epidemic and the degree of reduction of the infant mortality rate (IMR), from the time series approach.

The reason for the study of these two objectives is to determine the evolution of these variables over time and to verify if the obtained achievements have been equally reached by all the countries, or if, on the contrary, the progress has been unequal both between countries and within them. So that if these objectives have not been met, we can study the causes and possible explanations of this result, in order to shed some light on the decision making that helps these objectives can be achieved by 2030.

In order to be able to compare the analysis results of the different series, we have used the same methodology throughout the three chapters that make up the thesis: convergence analysis. The justification in the choice of the convergence analysis to carry out our research focuses on the fact that any policy aimed at either reducing infant mortality or ending the HIV / AIDS epidemic implies these variable should move go closer to their respective reference rates. Therefore, if all countries have reached the rates established in the MDGs /SDGs, this result could be interpreted as existence demonstration of convergence between countries. On the contrary, if only a part of the countries had been able to reach the reference rates in terms of infant mortality and prevalence of HIV / AIDS, then we could interpret the result as proof of the existence of a divergent behavior between countries. This fact would open the gate to the presence of different convergence clubs.

In this context, we consider recent advances in the field of convergence analysis, which are extremely useful to study whether the objectives proposed have been achieved and to identify those countries that have not managed to complete them. In this way, we are in a position to determine the forces that influence the creation of these convergence clubs by investigating the potentially explanatory variables previously collected in the literature.

This represents a radical change with respect to the type of data used mostly in the field of health economics, in which the use of time series data is strangely scarce. Against this, this thesis provides a different approach in the evaluation of health, from the perspective of time series and analysis of convergence, using macroeconomic data at national and international level, and through the application of novel econometric techniques not used until now in this type of studies.

Against this framework, this thesis tries to answer the following questions:

- Chapter 1: The recent evolution of the HIV / AIDS rates in Sub-Saharan Africa: Do differences still persist?

    - Have the MDGs / SDGs been met in terms of prevalence of HIV / AIDS in sub-Saharan Africa?
    - Are the differences in the evolution of HIV / AIDS maintained between sub-Saharan countries?
    - If the differences with regard to the expansion of HIV / AIDS in the countries of the African continent remain, what are the forces that lead to this result?
    - Are the explanatory variables collected in the literature as determinants of HIV / AIDS relevant in our analysis?

- Chapter 2: Do international child mortality rates converge?

    - Are the MDGs / SDGs met homogenously in terms of reducing child mortality?
    - Which countries have problems to achieve the MDGs / SDGs with respect to the objective established for the IMR?
    - What are the socioeconomic determinants that limit a country in the attempt to reduce its IMR?

- Chapter 3: Convergence in infant mortality rates: The Spanish case

    - Does convergence in the IMR exist between Spanish regions?

- Has the great economic recession influenced the regional differences of IMR in Spain?
- What have been the determining factors in the evolution of the IMR in Spain?

With the purpose of answer these questions, the Thesis compiles three scientific articles structured into chapters, which can be considered independently, although they are related to each other, prioritizing in this way the publication format.

Each of the chapters consists of two clearly differentiated parts. The first one is the convergence analysis, which includes the possibility of finding convergence clubs in the case of rejecting full convergence. The second one corresponds to the study of the explanatory factors that have led to the formation of these clubs in the case of having them.

The first chapter analyses the prevalence and incidence rates of HIV / AIDS in the countries of sub-Saharan Africa. Although the HIV epidemic is a global issue, the sub-Saharan region is the most affected region, with 25.6 million people living with HIV in 2016 and concentrating almost two-thirds of new infections. Countries in this region also face resource scarcity and depend on foreign aid to fight against HIV/AIDS. Accordingly, we find sufficient reason to geographically concentrate our research in the African continent.

In this chapter, after a preliminary study of the recent evolution of HIV prevalence and incidence rates, we proceed to test the null hypothesis of convergence for the entire sample, with the aim of verifying the existence of a unique pattern of behavior in the evolution of the analyzed rates. If we can reject the convergence hypothesis, and therefore the MDGs / SDGs have not been reached in the sub-Saharan region, we try to determine the presence of groups of countries with similar profiles trends in prevalence or incidence rates, or “convergence clubs”. Once the presence of these clubs is determined, we analyze their characteristics and explain the factors that make these different behaviors coexist.

The second chapter studies the reduction of international infant mortality rates (IMRs), which coincides with the fourth objective of the MDGs / SDGs. The IMR is considered one of the best global indicators to assess the welfare and health of a society, and this is the reason why this work has focused its attention on this MDG / SDG objective. Despite the great achievements made in reducing the IMR around the world, we cannot forget that around 15,000 children died every day in 2016 before reaching their fifth birthday, and more than five million children continue to die each year nowadays.

The procedure followed in this chapter is similar to the previous one. First, we evaluated the existence of a single pattern of behavior for the two samples used (Infant mortality rate of children under one year of age and infant mortality rate for children under five years of age). If we reject this hypothesis, we should analyze the existence different groups of convergence between countries. In the second phase of the study, we proceed to determine the forces that lead to the creation of these clubs.

To conclude, the third chapter of this Thesis explores the long-term evolution of infant mortality rates in the case of Spain. The justification in the choice of this essay comes from the need to compare the results obtained in the previous chapter, at the international level, in a context where different national health systems do not interfere in the results. In this line, it would not be strange to consider that regional IMR could provide more evidence in favor of the convergence hypothesis, as shown by Serenius et al. (2014) and Marlow (2014) in Sweden, Chen et al. (2016) in Switzerland, Sidebotham et al. (2014) in the England and Wales. However, this does not seem to be the case in Spain according to the seminal analyzes carried out by Regidor et al. (1994, 1995) and Regidor et al. (2002).

The application of the convergence analysis methodology, which acts as the backbone of this Thesis, allows us to exploit in greater depth the possible existence of stochastic convergence between the IMR of the Spanish regions. This offers us the opportunity to study the determinants that lead to obtaining the results in the previous analysis, as has been done in the preceding chapters.

*"La salud es un impulsor, un indicador y un resultado del desarrollo sostenible. Las personas sanas están mejor capacitadas para contribuir al desarrollo de sus países."*

Programa de las Naciones Unidas para el Desarrollo (PNUD)

# Introducción

La salud es uno de los pilares fundamentales para el desarrollo económico y del estado de bienestar de cualquier sociedad, por ello los principales organismos internacionales (ONU, Banco Mundial, Fondo Monetario Internacional, … etc.) han resaltado desde sus comienzos la importancia de la misma con numerosos proyectos a nivel mundial y la creación de instituciones internacionales dedicadas exclusivamente al campo de la salud, como la constitución de la Organización Mundial de la Salud (OMS) en 1948. Un claro ejemplo de ello es la Declaración del Milenio de las Naciones Unidas, el proyecto más ambicioso de esta organización hasta el momento, firmada en septiembre del año 2000 por 189 países miembros, en la cual se comprometían a perseguir ocho Objetivos de Desarrollo del Milenio (ODM) para el año 2015, siendo estos reemplazados en 2016 por los Objetivos del Desarrollo Sostenible (ODS).

Aunque todos los objetivos establecidos en los ODM tienen una relación directa o indirecta con la salud, hay tres de ellos dedicados exclusivamente a la misma: ODM4- Reducir la mortalidad infantil, ODM5- Mejorar la salud maternal y ODM6- Combatir el Virus de Inmunodeficiencia Humana VIH / Síndrome de Inmunodeficiencia Adquirida SIDA, la malaria y otras enfermedades. En la actualidad estos objetivos han sido englobados en el objetivo 3 de los ODS, denominado "Salud y bienestar", que propone para 2030: acabar con las muertes evitables de recién nacidos y de niños menores de 5 años, logrando que todos los países intenten reducir la mortalidad neonatal al menos hasta 12 por cada 1.000 nacidos vivos y la mortalidad de niños menores de 5 años al menos hasta 25 por cada 1.000 nacidos vivos. Poner fin a las epidemias del HIV/SIDA, la tuberculosis, la malaria y las enfermedades tropicales desatendidas. Así como combatir la hepatitis, las enfermedades transmitidas por el agua y otras enfermedades transmisibles.

Dado el interés que suscita el campo de la salud en la economía no es de extrañar que este campo haya atraído a un gran número de investigadores. Siguiendo el mismo interés parece relevante analizar cómo y en qué grado se han conseguido dos de los objetivos establecidos por las Naciones Unidas en los ODM. Concretamente, analizamos la evolución de la epidemia del VIH/SIDA y el grado de reducción de la tasa de mortalidad infantil (TMI), desde el enfoque de las series temporales.

La razón para el estudio de estos dos objetivos no es otro que determinar cuál ha sido la evolución de estas variables a lo largo del tiempo y comprobar si los logros obtenidos han sido igualmente alcanzados por todos los países, o si por el contrario, el progreso ha sido desigual tanto entre países como dentro de ellos. De forma que, si estos objetivos no se han visto cumplidos, podamos estudiar las causas y posibles explicaciones de este resultado, con el fin de arrojar algo de luz en la toma de decisiones que ayuden a que estos objetivos se cumplan para 2030.

Con el fin de poder comparar los resultados del análisis de las distintas series hemos utilizado la misma metodología a lo largo de los tres capítulos que componen la tesis: el análisis de convergencia.

La justificación en la elección del análisis de convergencia para llevar a cabo nuestra investigación se centra en el hecho de que cualquier política orientada, ya sea a reducir la mortalidad infantil o a acabar con la epidemia del VIH/SIDA, lleva implicada una disminución de la TMI en el caso de la primera, y la disminución en las tasas de prevalencia del VIH/SIDA en el caso de la segunda, hacia sus respectivas tasas de referencia. Por tanto, si todos los países hubiesen alcanzado las tasas establecidas en los ODM/ODS, este resultado podría interpretarse como una demostración de la existencia de convergencia entre los países. Por el contrario, si solo una parte de los países no hubiesen sido capaces de alcanzar las tasas de referencia en cuanto a mortalidad infantil y prevalencia del VIH/SIDA, entonces podríamos interpretar el resultado como prueba de la existencia un comportamiento divergente entre países, lo cual abre las puertas a la posibilidad de que existan diferentes clubs de convergencia dentro de la muestra.

En este contexto, consideramos que los recientes avances en el campo del análisis de convergencia son extremadamente útiles para estudiar si se han alcanzado los objetivos propuestos e identificar a aquellos países que no han conseguido completarlos. De esta forma, estamos en condiciones de, mediante la investigación de las variables potencialmente explicativas previamente recogidas en la literatura, determinar cuáles son las fuerzas que influyen en la creación de estos clubs de convergencia.

Esto supone un cambio radical con respecto al tipo de datos que se usan mayoritariamente en el campo de la economía de la salud, en el que el empleo de datos de series temporales es extrañamente escaso. Frente a ello, esta tesis aporta un enfoque distinto en la evaluación de la salud, desde la perspectiva del estudio de las series temporales y el análisis de convergencia, con datos macroeconómicos de ámbito nacional e internacional, y mediante la aplicación de novedosas técnicas econométricas no utilizadas hasta el momento en este tipo de estudios.

Dentro de este marco, los capítulos 1, 2 y 3 presentados en la Tesis tratan de responder a las siguientes preguntas:

- Capítulo 1: La reciente evolución de las tasas de VIH/SIDA en África Subsahariana: ¿Persisten las diferencias?

    - ¿Se han cumplido los ODM/ODS en cuanto a prevalencia del VIH/SIDA en África Subsahariana?
    - ¿Se mantienen las diferencias entre los distintos países de la región subsahariana en la evolución de la incidencia del HIV/SIDA?
    - Si se mantienen las diferencias respecto a la expansión del VIH/SIDA en los países del continente Africano, ¿Cuáles son las fuerzas que llevan a este resultado?
    - ¿Tienen relevancia las variables explicativas recogidas en la literatura como factores determinantes en el VIH/SIDA en nuestro análisis?

- Capítulo 2: ¿Existe convergencia en las tasas de mortalidad infantil internacionales?

  - ¿Se cumplen homogéneamente los ODM/ODS en cuanto a la reducción de la mortalidad infantil a nivel internacional?
  - ¿Qué países tienen problemas para cumplir los ODM/ODS respecto a objetivo establecido para la TMI?
  - ¿Cuáles son los determinantes socioeconómicos que limitan a un país en el intento de disminuir su TMI?

- Capítulo 3: Convergencia en tasa de mortalidad infantil: El caso de España

  - ¿Existe convergencia de las TMI entre las comunidades autónomas españolas?
  - ¿Ha influido la "Gran recesión" económica en las diferencias regionales de la TMI en España?
  - ¿Cuáles han sido los factores determinantes en la evolución de la TMI en España?

Con el fin de responder a estas cuestiones la presente Tesis compila tres artículos científicos estructurados en capítulos, que pueden considerarse de forma independiente, aunque estén relacionados entre sí, priorizando de esta manera el formato de publicación.

Cada uno de los capítulos consta de dos partes claramente diferenciadas: la primera es el análisis de convergencia para la muestra total y la identificación de clubs de convergencia en el caso de rechazar la anterior, la segunda parte corresponde al estudio de los factores explicativos que han dado lugar a la formación de estos clubs en el caso de haberlos.

En el primer capítulo se analizan las tasas de prevalencia e incidencia del VIH/SIDA en los países de África subsahariana. Aunque la epidemia del virus VIH es una asunto global, la región subsahariana es la más afectada, con 25.6 millones de personas viviendo con el VIH en 2016 y concentrando casi las dos terceras partes de las nuevas infecciones por el virus. En consecuencia, y si sumamos el problema de la escasez de recursos con la que se encuentran la mayoría de los países de esta región para hacer frente a la lucha contra el VIH/SIDA, encontramos razón suficiente para concentrar geográficamente nuestra investigación en el continente africano.

En este capítulo, tras un primer estudio de la reciente evolución de tasas de prevalencia e incidencia del VIH, procedemos a contrastar la hipótesis nula de convergencia para todo la muestra, con el objetivo de verificar la existencia de un único patrón de comportamiento en la evolución de las tasas en análisis. Si podemos rechazar la hipótesis de convergencia, y por lo tanto no se han alcanzado los ODM/ODS en la región subsahariana, intentamos determinar la presencia de "clubs de convergencia" (grupos de países con perfiles de tendencias similares de las tasas de prevalencia o incidencia). Una vez determinada la presencia de estos clubs, analizamos sus características y explicamos cuales son los factores que hacen que estos diferentes comportamientos coexistan.

El segundo capítulo estudia el cuarto objetivo de los ODM/ODS; la reducción en las tasas internacionales de mortalidad infantil. La tasa de mortalidad infantil es considerada uno de los mejores indicadores a nivel global para evaluar el bienestar y la salud de una sociedad, es por ello por lo que este trabajo ha centrado su atención en este objetivo de los ODM/ODS. A pesar de los grandes logros alcanzados en cuanto a reducción de las tasa de mortalidad infantil alrededor del mundo, no podemos olvidar que alrededor de 15,000 niños morían al día en 2016 antes de alcanzar su quinto cumpleaños, y más cinco millones de niños siguen muriendo cada año en la actualidad.

El procedimiento seguido en este capítulo es similar al anterior; en primer lugar evaluamos la existencia de un único patrón de comportamiento para las dos muestras utilizadas (tasa de mortalidad infantil de niños menores de un año y tasa de mortalidad infantil para niños menores de cinco años), o si por el contrario, existen diferentes grupos de convergencia entre países. En la segunda fase del estudio procedemos a determinar las fuerzas que llevan a la creación de estos clubs, en el caso de rechazar la hipótesis nula de convergencia para el total de la muestra.

Para finalizar, el capítulo tercero de esta Tesis indaga en la evolución a largo plazo de las tasas de mortalidad infantil en el caso de España. La justificación en la elección de este ensayo viene de la mano de la necesidad de comparar los resultados obtenidos en el capítulo anterior, en el ámbito internacional, en un contexto donde diferentes sistemas nacionales de salud no interfieran en los resultados. En esta línea, no sería extraño considerar que las TMI regional podrían proporcionar una mayor evidencia a favor de la hipótesis de convergencia, como demuestran Serenius et al. (2014) y Marlow (2014) en Suecia, Chen et al. (2016) en Suiza, Sidebotham et al. (2014) en Inglaterra y Gales. Sin embargo este no parece ser el caso teniendo en cuenta los análisis seminales en España llevados a cabo por Regidor et al. (1994, 1995) y Regidor et al. (2002). La aplicación de la metodología del análisis de convergencia, que ejerce de columna vertebral de esta Tesis, nos permite explotar en mayor profundidad la posible existencia de convergencia estocástica entre las TMI de las regiones españolas, proporcionándonos además la oportunidad de estudiar los determinantes que conducen a la obtención de los resultados en el análisis anterior, como se ha hecho en los capítulos precedentes.

# Chapter 1

# The recent evolution of the HIV rates in Sub-Saharan Africa: Do differences still persist?

## 1.1 Introduction

According to United Nations Development Programme (2005), the human immunodeficiency virus (HIV)/acquired immunodeficiency syndrome (AIDS) epidemic is the disease that has inflicted the single greatest reversal in human development in modern history. It has claimed more than 35 million lives so far, with 1.0 million people dying from HIV-related causes in 2016. Whilst this is a global problem, we should note that Sub-Saharan Africa is the most affected region, with 25.6 million people living with HIV in 2016, while it accounts for almost two thirds of the global total of new HIV infections. The situation was so complicated at the end of the XX century, with some prevalence rates close to 30% and incidence rates above 5%, that the

United Nations specifically included the combat against the spread of HIV as target 6 of its Millennium Development Goals (MDGs) in the Millennium Declaration. This document was signed in September 2000, with the aim of halting and beginning to reverse the spread of HIV by 2015. These MDGs were subsequently substituted by the Sustainable Development Goals (SDGs) and, again, the reduction of the spread of HIV/AIDS explicitly appears in the set of 17 goals included the 2030 Development Agenda signed by the 193 countries of the United Nations General Assembly on 25 September 2015.

The application of the different United Nations' HIV-oriented policies has been relatively successful and there is a general consensus that they are closely related to the decline of the spread of HIV/AIDS during the XXI century. New HIV infections fell by 39% during 2000-2016 and the development of new effective antiretroviral drugs have helped to control the virus and prevent the transmission of the virus. The consequences are a remarkable reduction in HIV-related deaths[1] and a more healthy life for HIV-infected people.

However, this does not mean that the battle against this epidemic is over. We can still appreciate high values of HIV prevalence and, to lesser extent, incidence rates in some Sub-Saharan countries. This leads us to consider that the effect of the different international and local policies has not had the same effect in all the Sub-Saharan countries, although they share similar social and economic conditions. It also suggest the possible existence of different patterns of behavior, which could be caused by the existence of zones where HIV transmission is not controlled and where, as a consequence, more resources and efforts should be dedicated to control it.

[1] According to the World Health Organization (WHO) data, the use of the antiretroviral therapy (ART) has saved the lives of 13.1 million people during 2000-2016.

We can interpret the existence of these different patterns of behavior from a convergence perspective. If the values of the largest/smallest prevalence/incidence rates move towards the smallest/largest, there is a single pattern of behavior between all of them. By contrast, if the rates diverge, then multiple patterns of behaviors can be found. In our context, this result would imply a lack of control of the epidemic and, as a consequence, that the HIV-oriented policies have not been successful as was expected according to the MDG and SDG agendas.

In this context, the paper tries to find answer to the following questions:

- Have the MDGs / SDGs been met in terms of prevalence of HIV / AIDS in sub-Saharan Africa?
- Are the differences in the evolution of HIV / AIDS maintained between sub-Saharan countries?
- If the differences with regard to the expansion of HIV / AIDS in the countries of the African continent remain, what are the forces that lead to this result?
- Are the explanatory variables collected in the literature as determinants of HIV / AIDS relevant in our analysis?

In order to answer the previous questions the purpose of the present paper is to analyze the null hypothesis of convergence for the above mentioned prevalence/incidence HIV rates. If we cannot reject this null hypothesis, it means that there is a single pattern of behavior for these rates across the Sub-Saharan countries. In the case of convergence toward the low rates we could considered the HIV epidemic as controlled, although not eradicated. By contrast, if we can reject the null hypothesis of convergence, then multiple patterns of behaviors exist and, consequently, we should conclude that more efforts are necessary to reduce the HIV rates before we can consider this epidemic to be controlled.

The rest of the paper is organized as follows. We first describe the database employed in section 1.2, paying attention to the recent evolution of the two measures that we will employ in the paper, namely, the prevalence and the incidence rates. Later, in section 1.3 we test for the null hypothesis of convergence in order to verify the existence of a single pattern of behavior in the evolution of these rates. If we reject this hypothesis, we will try to determine the possible presence of convergence clubs or groups of countries with similar trend profiles of prevalence or incidence rates. If these clubs are found, in section 1.4 we will analyze their characteristics and, additionally, explain why these different behaviors coexist. In particular, the role played by female education may help us to explain them and, furthermore, to establish some policies that could help to fight the HIV/AIDS epidemic in Sub-Saharan countries. The paper ends with a review of the most important conclusions.

### 1.2. Database description

In order to assess the evolution of HIV infection in the Sub-Saharan countries, we will use two different indicators, namely, the prevalence and the incidence of HIV. Incidence is defined as the number of new HIV infections among uninfected populations aged 15-49, expressed per 100 uninfected population in the year before the period, whilst prevalence refers to the percentage of people aged 15-49 who are infected with HIV. We also disaggregate prevalence by gender. We have selected the countries that have full data for 1990-2016, the data having been collected from the World Health Organization database.

Table 1.1 reports the values of the female/male prevalence (FPHIV/MPHIV) and the total incidence (TIHIV) rates for the Sub-Saharan countries for different years. We can observe a great heterogeneity in the data. For instance, considering the data of 1990, Comoros, Gambia, Guinea-Bissau, Madagascar, Mauritania, Somalia and South Sudan showed female prevalence (FPHIV) levels close to 0, whilst the rates of Zimbabwe and Uganda were around 15 percent. If we take into account the data at the end of the sample, we can observe that half of the countries included in the sample exhibit a female prevalence rate lower than (or equal to) 1 percent, whilst Swaziland, Lesotho, South Africa and Botswana present rates greater than 10 percent. The highest recorded value is 32.1 percent, corresponding to Swaziland in 1999. If we now consider the male prevalence rates (MPHIV), we can first observe that the values are generally lower than those observed for the female case, with more than two thirds of the countries showing male prevalence rates lower than 1 percent. The highest historic value is 10.8 percent and it again corresponds to Swaziland, but in 1997. The countries with the highest male prevalence rates at the end of the sample are Lesotho (6.3 percent), Botswana (5.4 percent), Zambia (4.1 percent) and South Africa (4.0 percent). Finally, we should analyze the total incidence rates (TIHIV). The highest value (6.3 infected per 100 uninfected population) is recorded in Swaziland in 1996. We can also observe that 29 countries never show incidence rates greater than 1, whilst only Swaziland presents incidence rates greater than 1 for all the years considered. In 1990, the countries with the highest incidence were Swaziland (4.2 infected per 100 uninfected population), Lesotho (3.4 infected per 100 uninfected population), Botswana (3.2 infected per 100 uninfected population), South Africa (2.3 infected per 100 uninfected population), Zimbabwe (2.1 infected per 100 uninfected population), Namibia (2.1 infected per 100 uninfected population), Mozambique (2.0 infected per 100 uninfected population), Malawi (1.5 infected per 100 uninfected population) and Zambia (1.4 infected per 100 uninfected population) exhibiting incidence rates greater than 1, while

Lesotho (2.3 infected per 100 uninfected population) and Swaziland (1.7 infected per 100 uninfected population) presented the highest rates at the end of the sample.

A clear conclusion can be drawn from the previous descriptive analysis: HIV spread is far from being homogenous across the Sub-Saharan African countries. To support this idea, Figure 1.1 presents the evolution of the dispersion of the data, providing some useful initial insights about convergence for the different HIV measures considered. We can observe a very clear hump-shaped evolution in the standard deviation of the female prevalence rates, whose maximum value was in 1998. The dispersion remains relatively constant up to 2000, when a very clear decline begins. However, the dispersion value at the end of the data is still greater than the initial one. So, this figure leads us to cast some doubts about convergence, in spite of appreciating an evident effort to reduce the HIV female prevalence rates in the countries especially affected by this problem.

The evolution of the standard deviation of both the HIV male prevalence rates (MPHIV) and total incidence rates (TIHIV) is somewhat different. The curves also exhibit a humped shape, less markedly than that of the female prevalence rates. The maximum values are in 1997 and 1996, respectively, for male and total incidence rates, with these maximum values taking more modest values (2.7 and 1.6) if we compare them with the maximum value of the dispersion of the female prevalence rate (6.9). Finally, if we compare the initial and the final dispersion values, we observe that the MPHIV rates hardly vary, whilst the TIHIV rates reduce from 1.0 to 0.5. So, we have some evidence in favor of the convergence hypothesis, in the sense that the countries with the highest MPHIV and TIHIV rates at the beginning of the sample have reduced these rates so a catching-up process may have occurred.

These results lead us to conclude that the TIHIV and MPHIV rates may exhibit convergence, which would involve the existence of a single pattern of behavior for these two variables. By contrast, the total and the female prevalence rates may diverge and, consequently, different patterns of behavior may co-exist.

In any case, these results are simply descriptive and, therefore, we should confirm this initial evidence by using more appropriate methods that directly test for the convergence null hypothesis. This is the aim of the next section.

### 1.3. Testing for convergence in HIV

The previous analysis has shown the disparities that exist between the evolution of the prevalence and the incidence HIV rates across Sub-Saharan countries, which offer some support to the existence of convergence and can be interpreted as a first and very important step towards the real possibility of ending this epidemic. To test for the null hypothesis of convergence we have followed the recent papers of Phillips and Sul (2007, 2009) (PS hereafter) where they develop a framework that allows us, first, to test for the convergence hypothesis and, if this hypothesis is rejected, to determine the composition of the different convergence clubs, if they exist. The PS methodology is closely related to the standard σ-convergence analysis, a concept introduced by Barro and Sala-i-Martin (1990) and Barro et al. (1991), in that it tests for the decline of the variable of interest over time in the cross-sectional dispersion. However, it clearly outperforms the classical convergence analysis, namely, the abovementioned σ-convergence concept and the β-convergence concept, which was first introduced by Baumol (1986). The PS methodology is based on a general nonlinear time-varying factor model, which admits the presence of transitional heterogeneity. Additionally, it is flexible with respect to the time properties of the variables under analysis because it does not impose any particular assumption about them. Finally, it is clearly free of the criticism received by the β-convergence analysis in De Long (1988) and Quah (1993). As a consequence, the use of this methodology has recently become very popular in convergence analysis.

Following these authors, let us consider that $X_{it}$ represents the different HIV measures considered in this paper, with i=1, 2, … 45 (the 45 sub-Saharan countries) and t is the year in the period 1990-2016. This variable can be decomposed as $X_{it} = \delta_{it}\,\mu_t$, where $\mu_t$ and $\delta_{it}$ are the common and the idiosyncratic components, respectively. PS suggest testing for convergence by analyzing whether $\delta_{it}$ converges towards δ. To do so, they first define the relative transition component or transition coefficient:

$$h_{it} = \frac{X_{it}}{N^{-1}\sum_{i=1}^{N} X_{it}} = \frac{\delta_{it}}{N^{-1}\sum_{i=1}^{N}\delta_{it}} \tag{1.1}$$

In the presence of convergence, $h_{it}$ should converge towards unity, while its cross-sectional variance, $H_{it}$, is defined as follows:

$$H_{it} = N^{-1}\sum_{i=1}^{N}(h_{it}-1)^2 \tag{1.2}$$

and should go to 0 when T goes towards infinity. Then, PS test for convergence by estimating the following equation:

$$log\frac{H_1}{H_t} - 2log[\log(t)] = \alpha + \beta\log(t) + u_t,\ t = T_o, \dots, T \tag{1.3}$$

with $T_o = [rT]$, and r=0.3, as suggested by PS. Equation (1.3) is commonly known as the log-t regression. The null hypothesis of convergence is rejected whenever parameter β is lower than 0. PS suggest estimating model (1.3) by methods which correct for the presence of autocorrelation and heteroscedasticity and, later, employ the t-statistic to test the null hypothesis β=0. The use of these robust methods ensures that this t-ratio converges towards a standard N(0,1) distribution and, therefore, we will reject the null hypothesis of convergence whenever this t-statistic takes values lower than -1.65.

If we reject convergence, PS propose the following robust clustering algorithm for identifying clubs in a panel of countries:

i. Order the N countries according to their final values.

ii. Starting from the highest-order countries, add adjacent countries from our ordered list and estimate model (1.3). Then, select the core group by maximizing the value of the convergence t-statistic, subject to the restriction that it is greater than -1.65.

iii. Continue adding one country at a time of the remaining countries to the core group, and re-estimate model (1.3) for each formation. Use the sign criterion (t-statistic >0) to decide whether a country should join the core group.

iv. For the remaining countries, repeat steps (ii)–(iii) iteratively and stop when clubs can no longer be formed. If the last group does not have a convergence pattern, conclude that its members diverge.

PS recommend performing club merging tests after running the algorithm using equation (1.3) in order to avoid the over-estimation of the number of clubs.

Finally, we have followed the suggestion of PS and we have extracted the trend components of the series by filtering them using the Hodrick and Prescott (1997) filter, applying the standard value λ=400.

The results that we have obtained are presented in Table 1.2. This table reports the results for the female HIV prevalence rates (FPHIV) and male HIV prevalence rates (MPHIV), as well as for the total HIV incidence rates (TIHIV). The results obtained are mixed. We can first observe that the convergence null hypothesis is not rejected for either the TIHIV or the MPHIV, as was suggested by the analysis of Figure 1.1. This result could be interpreted as a successful implantation of the health policies in the recent years, especially those related to the MDGs and, to lesser extent, the SDGs. In order to explore the influence of these two agendas on HIV evolution, we have additionally analyzed the null hypothesis of convergence for several periods of time. As we can observe, the MPHIV quickly converge, given that we cannot reject the convergence null hypothesis since 2005, whilst the convergence of the TIHIV occurred later. Thus, we can affirm that it took longer to control the evolution of the TIHIV rates than the evolution of the MPHIV rates. Similarly, the non-rejection of the convergence null hypothesis implies that the two rates show a unique pattern of behavior across the countries included in our sample. Consequently, it could be considered that the measures recommended in both MDG and SDG agendas have been useful to control the HIV epidemic, at least as far as the MPHIV and the TIHIV rates are concerned.

If we now analyze the behavior of the FPHIV rates, the results reported in Table 1.2 show that we can clearly reject the convergence null hypothesis for this rate. It is true that the evidence against this hypothesis is now smaller than it was in 2004, but it still is large enough to allow us to reject the presence of a single pattern of behavior among the Sub-Saharan countries. This implies that more efforts are necessary to control the evolution of the FPHIV rates, making it very difficult to adopt similar policies in all the countries. Rather, it would seem to be sensible to consider adopting country-specific policies or, at least, different policies for some groups of countries, if we can group some of them by the similar evolution of their FPHIV rates. In this regard, we should note that, once we have found divergence, it is advisable to analyze whether convergence clubs exist, showing the presence of groups of countries whose FPHIV rates share a similar pattern of behavior. To that end, we can employ the PS algorithm. The results that we have obtained from the use of the cluster algorithm proposed by these authors, lead us to conclude in favor of the existence of two clubs with statistically different behaviors. Figure 1.2 maps these clubs in different colors. The first club, which will be referred to as club 1, is composed of the following countries: Botswana, Equatorial Guinea, Lesotho, Malawi, Mozambique, Namibia, Sierra Leone, South Africa, Swaziland, Uganda, Zimbabwe and Zambia, whilst the rest of the countries are included in the second club, which will be referred to as club 2. This result is quite understandable given that club 1 includes the 10 countries with the highest female prevalence rate in 2016 plus Equatorial Guinea, the 12$^{th}$ country in 2016 and with a relatively important growth in the recent years, and Sierra Leone, with a small prevalence rate but showing a final value very close to its maximum, which was in 2005. Additionally, Equatorial Guinea and Sierra Leone are the only two countries that have increased their FPHIV rates since 2000. Behaviors mentioned above are reflected in the transition coefficient evolution of club 1 and club 2 shown in Figure 1.3. In any event, their inclusion in club 1 should be interpreted with some caution, given that these results are reflecting the relatively bad behavior of the female prevalence rates in recent years, which could originate a problem in the future, more than the current situation.

In order to reflect the evolution of the two groups, Figure 1.4 represent the evolution of the average values of the FPHIV rates of the countries in the estimated clubs. Once again, we can appreciate a hump-shaped curve for both of them, although a much sharper one for the countries included in club 1. We can also see that the initial distance between the two curves is around 4 percentage points. It grew to 11 points in 1999, when it attains its maximum, and has decreased since then to 6.2 points in 2016, the minimum distance since 1992. So, our data again reflects the positive effects of the UN recommendations and the local policies to control the evolution of the FPHIV rates. However, the important differences that emerged at the end of the last century have not disappeared yet and, as a consequence, more efforts are still necessary, first, to move the club 1 FPHIV rates towards those of club 2 and, second, to reduce all of them towards 0, finally eradicating this painful pandemic. To that end, the use of appropriate policies is necessary. The next section is devoted to comparing the evolution of the countries in clubs 1 and 2 with the aim of identifying which polices may offer better results in order, first, to control the evolution of the FPHIV rates and, second, to be in a position of winning the battle against HIV/AIDS in the near future.

### 1.4. What forces drive the creation of the clubs?

The results of the previous section prove that the evolution of the FPHIV rates in Sub-Saharan Africa exhibits different patterns of behavior and that it is possible to split the countries into two separate groups. Now we need to explain why these differences occur, which could help us to draw some insights to know which policies are more relevant in the fight against this pandemic. The identification of the key policies is not an easy task, especially if we recognize the existence of multiple factors that may explain these differences. We can find some of them in the literature. For instance, some authors have recently analyzed the role of education in the recent decline of the HIV prevalence rates. In this regard, we can cite the papers of Alsan and Cutler (2013), Behrman (2015) and De Neve et al. (2015), who study the cases of Uganda, Malawi and Botswana, respectively. Following De Neve et al. (2015), we can appreciate that increasing the number of years of secondary schooling had a large protective effect against HIV risk in Botswana, particularly for women. Similar results are obtained for the previously mentioned authors. Therefore, we can understand education as a social vaccine to reduce the spread of HIV, an idea supported by authors such as Vandemoortele and Delamonica (2000) and Amaugo et al. (2014).

We should also consider the evolution of the economy as a very important factor to battle against HIV/AIDS. Some authors, such as Mbirimtengerenji (2007), consider that poverty, and its consequences, is a crucial point of HIV transmission. It comes as no surprise that HIV/AIDS is often referred to as a disease of poverty. However, as Mufune (2014) points out, the relationship between poverty and HIV transmission is not clear, because countries with similar levels of poverty do not exhibit similar levels of prevalence rate. Thus, some economic variables need to be included to verify whether this relationship exists and, if so, its direction.

From a different point of view, some authors have analyzed the influence of the use of contraceptive methods on HIV transmission. We can cite the recent papers of Magadi and Magadi (2017) and Mabaso et al. (2018), where the Kenyan and the South-African cases are analyzed, respectively. These authors find some relationship between the use of the condom as a contraceptive method and HIV transmission. However, the direction of this relationship is not clear and an appropriate use seems to be more crucial to diminish the risk of HIV transmission than the mere use of this method. In any event, we consider it interesting to study whether this effect is valid for other countries in order to find a general result.

Finally, we consider the degree of urbanization of the country. Since the paper by Dyson (2003), some authors have studied whether HIV/AIDS presence is more an urban than a rural problem. We should also cite the papers of Asiedu (2012), Hajizadeh et al. (2014) and Taaffe et al. (2016) in this regard. According to them, HIV/AIDS seems to be more related to urban population, although the rural/urban differences have diminished in the recent years.

Having selected the different factors, we should now choose the most appropriate variables. In spite of the problems of data availability, we have been able to collect a relatively high number of explanatory variables, although we have had to remove some countries from the study due to missing data, finally using 36 of the 45 countries that were initially employed for the convergence analysis. The list of explanatory variables and the removed countries are reported in Appendix A. Table 1.3 presents the mean of the most relevant variables for each club. As we can see, the youth unemployment rates and female condom use take very different values for each club, whilst the differences for the other variables are lower. We can also appreciate that per capita Gross Domestic Product (GDP), the Human Development Index (HDI) and female condom use take larger values for club 1 than for club 2. This result could be considered somewhat counter-intuitive because we would expect the countries with the highest FPHIV rates (club 1) to show smaller values of per capita GDP and the Human Development Index and a lower percentage of females that use condoms in their sexual relationships than the countries with the smallest FPHIV rates.

However, this is a mere descriptive analysis and it would be more appropriate to use more robust methods to determine which variables allow us to explain why a particular country is assigned to club 1 or 2. To that end, we have estimated a probit model. Given that club 2 is the largest, the dependent variable takes the value 0 if the country has been included in club 2 and 1 when the country belongs to club 1. The results we have obtained are shown in Table 1.4. These results have been obtained by using a general-to-particular encompassing strategy. We have first estimated a general model by ordinary least squares with all the variables included in Table A1 and, then, removed the non-significant variables one-by-one. Then, we have estimated the probit model with the selected variables. To avoid the problems generated by the possible presence of non-observed heterogeneity, we have employed robust estimation methods following White (1980). As can be seen, we can discriminate the characteristics of the two clubs by employing just four variables. Countries with a high percentage of female population aged 15-24 who used a condom during their last intercourse in the last 12 months are mostly included in club 2. This result is somewhat counterintuitive, as we have mentioned earlier, and should be interpreted with some caution or it could lead us to draw erroneous conclusions. First, we should recall the results reflected in Table 1.3., where countries included in club 1 exhibit the highest percentages of female condom use. However, we should not conclude that this is the cause of the high female prevalence rates in these countries. Rather, these large percentages in the use of condom are denoting the impact of health policies which have extended the use of condom and, as a consequence, these countries have controlled the incidence rates, the first step to stop the extension of the HIV epidemic. Furthermore, the results of Magadi and Magadi (2017) and Mabaso et al. (2018) had previously warned of the importance of a consistent use of the condom, more than its mere use. So, more efforts seem to be necessary in this regard.

Next variable included in the model is female persistence to the last year of primary, which reflects the percentage of females enrolled in the first year of primary school who eventually reach the last year of primary education. So, if the female persistence value is high, large female prevalence rates are less likely. Moreover, if we calculate the marginal effects, we can observe that, if the countries included in club 1 could increase their female persistence to the levels of the countries in club 2, the probability of being included in club 1 would be reduced by 12%. These data prove the importance of education in the fight against the HIV/AIDS pandemic. This result is absolutely related to the abovementioned literature and offers additional evidence on the importance of education in the fight against the HIV epidemic. However, our results are more general that the previous studies because our evidence is not valid for a single country, but rather for the set of countries with large female prevalence rates.

The third variable is the percentage of urban population. As we can see, the countries included in club 2 seem to exhibit highest values of this variable. Thus, it seems that the advantages of living in relatively large cities are more important than its disadvantages in order to control the HIV/AIDS epidemic.

Finally, the youth unemployment rate reflects the evolution of the economy in the sense that it is expected that countries with a higher HDI value and a low youth employment rate should exhibit a lower FPHIV rates than the rest.

In order to appreciate the relationship between the explanatory variables included in the model and the probability of a particular country being assigned to club 1, we have graphically represented this probability for several values of female persistence to the last year of primary (FPP) and youth male unemployment rates (YMUR), whilst the rest of the variables take the average values of the countries included in club 1. Figure 1.5 reflects the evolution of the probability when FPP takes values in the (30, 90) interval. As we can observe, the probability is greater than 0.9 whenever FPP is lower than 55%, whilst it is 0.8 if FPP takes greater values than the mean of the total sample (63%). Similarly, the probability of being included in club 1 is not lower than 50% until FPP takes values greater than 77%. If we take into account that only six countries exhibit these values in our sample, we can understand how difficult it is to move from club 1 to club 2 for a particular country only using educational policies. Similar results are obtained if we analyze Figure 1.6. We can observe in this case that values of the YMUR greater than 30% imply probabilities greater than 90% of being included in club 1. If the YMUR take the total sample median value (15.4%), this probability reduces to 62% and it is lower than 50% when the YMUR takes the total sample mean (10%). Finally, if we combine the evolution of both variables, the results are more encouraging. For instance, if the policies adopted by a country included in club 1 lead it to increase the FPP and reduce YMUR to the total sample median values, then the probability of being included in club 1 is just 42%. So, the combination of socioeconomics variables, together with maintaining the current health policies, may help to definitively control the HIV/AIDS epidemic in the near future.

### 1.5. Conclusions

This paper analyses the evolution of the HIV/AIDS epidemic in Sub-Saharan Africa from a time series perspective. We have focused on two different measures of this problem: the HIV prevalence rate and the HIV incidence rate. Given the data availability, we have disaggregated the HIV prevalence rate by gender. The use of the recent advances in convergence analysis has allowed us to study whether these two measures converge and exhibit a unique behavior across the sample size, which covers 1990-2016, or whether, by contrast, they exhibit different patterns of behavior. If the latter case is found, then we should conclude that the epidemic is not controlled and that some countries still exhibit a different pattern of behavior to the rest, in spite of the great effort made by the Sub-Saharan countries with extraordinary support from the international community.

Our results lead us to conclude that the male HIV prevalence and the total HIV incidence rates show a unique pattern of behavior. This implies that those countries with the largest rates at the beginning of the sample have reduced the distance with respect to those with the smallest rates by decreasing their prevalence/incidence rates, which could be interpreted as the success of the policies adopted. Moreover, we have verified that this result is relatively recent, in the sense that a different conclusion is drawn when the 1990-2010 sample is considered, at least for the total HIV incidence rate. Consequently, it seems that all the policies adopted, which were defined by the MDG agenda, have been quite effective in the fight against the HIV/AIDS epidemic and both these rates; MPHIV and TIHIV, seem to be under control.

However, we can reject the presence of a unique pattern of behavior for the female HIV prevalence rate. Rather, we observe that the female rate of the countries with the highest prevalence rate (Botswana, Lesotho, Malawi, Mozambique, Namibia, South Africa, Swaziland, Uganda, Zimbabwe and Zambia) plus Equatorial Guinea and Sierra Leone, with a somewhat large growth in the recent years, behaves differently with respect to the rest of the Sub-Saharan countries. Thus, this convergence analysis helps us to both identify the existence of a group of countries where the HIV/AIDS is not controlled yet.

Having found the origin of the differences, we then tried to determine which factors may drive them. To that end, we have estimated a probit model with the dependent variable taking the value 0 if the country belongs to club 2 (the largest club) and 1 if the country belongs to club 1 (the club with the countries that exhibit largest female HIV prevalence rates). We have used several explanatory variables that capture the different socioeconomics aspects of these countries, mainly related to education, economics and sexual habits, all of them commonly identified as key factors to understand the

HIV/AIDS transmission. The variables finally included in the model are the percentage of urban population, the male youth unemployment rate, the percentage of the female population aged 15-24 who used a condom in their last intercourse in the last 12 months and the percentage of female children enrolled in the first year of primary school who eventually reach the last year of primary education. The results obtained imply that the probability of passing from club 1 to club 2 increases with an improvement of the economic situation (as proxied by the YMUR), the more consistent use of the contraceptive methods and, especially, the increment of the female education level. We have analyzed the evolution of the probability of being included in club 1 when the socioeconomic variables change. The combination of an increase in female persistence to the last year of primary and a reduction of the male youth unemployment rate to their respective sample mean would be quite effective and would reduce the probability of being assigned to club 1 to below 50%. This result reinforces the idea that general education improvements, especially those addressed to the female population, are crucial to eradicate the HIV/AIDS pandemic, although they should be accompanied by taking appropriate decisions on both economic and health policies.

Finally, we should note that the adoption of our suggested policies should be accompanied by the continuation of the support of the international community. We must recognize that the progress made in the fight against HIV/AIDS has been partially achieved due to a very large financial investment from the international community, as Oberth and Whiteside (2016) note. If this support is not maintained, it does not seem possible for these countries to provide sufficient resources to the HIV/AIDS fight and, consequently, all the achievements might suffer a serious setback.

## Table 1.1. Descriptive statistics

| Country Name | Country Code | FPHIV 1990 | FPHIV 2000 | FPHIV 2016 | MPHIV 1990 | MPHIV 2000 | MPHIV 2016 | TIHIV 1990 | TIHIV 2009 | TIHIV 2016 |
|---|---|---|---|---|---|---|---|---|---|---|
| Angola | AGO | 0.2 | 0.6 | 0.8 | 0.1 | 0.3 | 0.4 | 0.06 | 0.18 | 0.17 |
| Benin | BEN | 0.2 | 1.2 | 0.5 | 0.2 | 0.7 | 0.3 | 0.10 | 0.16 | 0.06 |
| Botswana | BWA | 7.9 | 22.5 | 10.2 | 3.4 | 8.6 | 5.4 | 2.63 | 3.20 | 0.93 |
| Burkina Faso | BFA | 4.6 | 1.7 | 0.6 | 1.7 | 0.6 | 0.4 | 0.54 | 0.07 | 0.03 |
| Burundi | BDI | 0.5 | 2.0 | 0.4 | 0.2 | 0.8 | 0.3 | 0.19 | 0.30 | 0.03 |
| Cabo Verde | CPV | 0.6 | 0.9 | 0.5 | 0.4 | 0.5 | 0.3 | 0.28 | 0.69 | 0.23 |
| Cameroon | CMR | 1.2 | 5.2 | 2.3 | 0.4 | 1.6 | 0.9 | 1.29 | 0.52 | 0.32 |
| Central African Republic | CAF | 2.4 | 4.0 | 1.9 | 1.7 | 2.0 | 1.2 | | | |
| Chad | TCD | 0.8 | 1.4 | 0.5 | 0.3 | 0.5 | 0.3 | 0.20 | 0.23 | 0.06 |
| Comoros | COM | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.01 | 0.01 | 0.01 |
| Congo, Dem. Rep. | COD | 0.5 | 1.3 | 0.3 | 1.4 | 0.8 | 0.1 | 0.26 | 0.20 | 0.03 |
| Congo, Rep. | COG | 2.2 | 1.5 | 1.4 | 1.8 | 0.8 | 0.8 | 0.92 | 0.27 | 0.28 |
| Cote d'Ivoire | CIV | 2.9 | 4.1 | 1.1 | 0.8 | 0.8 | 0.5 | 0.81 | 0.60 | 0.13 |
| Equatorial Guinea | GNQ | 1.3 | 1.8 | 2.6 | 0.7 | 0.9 | 1.4 | 0.26 | 0.55 | 0.49 |
| Eritrea | ERI | 0.3 | 1.2 | 0.2 | 0.1 | 0.4 | 0.2 | 0.13 | 0.15 | 0.02 |
| Ethiopia | ETH | 1.2 | 1.7 | 0.5 | 0.6 | 0.7 | 0.4 | 0.46 | 0.23 | 0.05 |
| Gabon | GAB | 1.8 | 7.6 | 2.0 | 0.3 | 1.4 | 0.6 | 0.35 | 0.80 | 0.14 |
| Gambia, The | GMB | 0.1 | 0.7 | 0.4 | 0.1 | 0.1 | 0.1 | 0.03 | 0.31 | 0.12 |
| Ghana | GHA | 1.1 | 2.5 | 1.0 | 0.8 | 0.7 | 0.4 | 0.55 | 0.25 | 0.12 |
| Guinea | GIN | 1.1 | 1.6 | 1.0 | 0.4 | 0.5 | 0.4 | 0.24 | 0.17 | 0.10 |
| Guinea-Bissau | GNB | 0.1 | 2.0 | 0.8 | 0.1 | 1.1 | 0.5 | 0.06 | 0.66 | 0.12 |
| Kenya | KEN | 7.4 | 8.1 | 3.5 | 3.7 | 3.3 | 1.9 | 1.64 | 0.64 | 0.25 |
| Lesotho | LSO | 1.4 | 18.9 | 13.9 | 0.8 | 7.7 | 6.3 | 0.73 | 3.42 | 2.27 |
| Liberia | LBR | 3.7 | 3.6 | 1.2 | 1.9 | 1.4 | 0.6 | 0.78 | 0.24 | 0.11 |
| Madagascar | MDG | 0.1 | 0.2 | 0.1 | 0.1 | 0.1 | 0.2 | 0.01 | 0.03 | 0.03 |
| Malawi | MWI | 8.4 | 10.3 | 4.5 | 2.7 | 3.0 | 2.2 | 1.99 | 1.54 | 0.41 |
| Mali | MLI | 0.7 | 1.5 | 0.6 | 0.9 | 0.8 | 0.3 | 0.19 | 0.18 | 0.05 |
| Mauritania | MRT | 0.1 | 0.5 | 0.1 | 0.1 | 0.8 | 0.2 | | | |
| Mozambique | MOZ | 1.1 | 6.8 | 4.6 | 1.0 | 4.1 | 2.8 | 0.52 | 1.96 | 0.66 |
| Namibia | NAM | 1.7 | 12.5 | 6.1 | 0.9 | 5.5 | 3.2 | 0.72 | 2.05 | 0.78 |
| Niger | NER | 0.3 | 0.9 | 0.2 | 0.2 | 0.7 | 0.1 | 0.09 | 0.13 | 0.02 |
| Nigeria | NGA | 1.6 | 3.3 | 1.6 | 1.1 | 2.3 | 1.0 | 0.29 | 0.50 | 0.20 |
| Rwanda | RWA | 2.2 | 3.7 | 1.3 | 0.8 | 0.9 | 0.7 | 0.60 | 0.50 | 0.13 |
| Senegal | SEN | 0.2 | 0.7 | 0.1 | 0.1 | 0.4 | 0.1 | 0.04 | 0.11 | 0.01 |
| Sierra Leone | SLE | 0.5 | 1.1 | 1.3 | 0.3 | 0.6 | 0.7 | | | |
| Somalia | SOM | 0.1 | 0.3 | 0.1 | 0.1 | 0.2 | 0.1 | 0.04 | 0.08 | 0.03 |
| South Africa | ZAF | 1.0 | 16.8 | 10.4 | 0.5 | 5.9 | 4.0 | 0.32 | 2.31 | 0.99 |
| South Sudan | SSD | 0.1 | 1.9 | 1.1 | 0.1 | 1.0 | 0.6 | 0.02 | 0.57 | 0.23 |
| Sudan | SDN | 0.1 | 0.1 | 0.2 | 0.1 | 0.1 | 0.1 | 0.01 | 0.02 | 0.02 |

| Country Name | Country Code | FPHIV 1990 | FPHIV 2000 | FPHIV 2016 | MPHIV 1990 | MPHIV 2000 | MPHIV 2016 | TIHIV 1990 | TIHIV 2009 | TIHIV 2016 |
|---|---|---|---|---|---|---|---|---|---|---|
| Swaziland | SWZ | 2.1 | 31.5 | 17.6 | 1.5 | 9.3 | 3.8 | 1.17 | 4.19 | 1.66 |
| Tanzania | TZA | 5.4 | 6.1 | 2.3 | 3.6 | 3.3 | 0.9 | 1.30 | 0.72 | 0.21 |
| Togo | TGO | 1.2 | 3.3 | 0.8 | 0.4 | 0.9 | 0.4 | 0.37 | 0.55 | 0.10 |
| Uganda | UGA | 13.5 | 5.1 | 3.8 | 5.3 | 2.1 | 1.9 | 1.69 | 0.72 | 0.30 |
| Zambia | ZMB | 8.8 | 11.9 | 6.9 | 5.1 | 6.1 | 4.1 | 2.34 | 1.38 | 0.73 |
| Zimbabwe | ZWE | 14.6 | 14.4 | 5.7 | 6.0 | 5.9 | 2.8 | 3.87 | 2.10 | 0.54 |

The first column of this table presents the countries included in the study, whilst the second column presents their corresponding acronyms.

Columns 3-5 present the values of the female HIV prevalence rates (percentage of females aged 15-49 who are infected with HIV) for 1990, 2000 and 2016, respectively.

Columns 6-8 present the values of the male HIV prevalence rates (percentage of males aged 15-49 who are infected with HIV) for 1990, 2000 and 2016, respectively.

Columns 9-11 present the values of the total HIV incidence rates (number of new HIV infections among uninfected populations aged 15-49, expressed per 100 uninfected population in the year before the period) for 1990, 2000 and 2016, respectively

Table 1.2. Testing for convergence

| Sample | FPHIV | MPHIV | TIHIV |
|---|---|---|---|
| 1990-2004 | -67.1* | -2.4* | -13.2* |
| 1990-2005 | -53.9* | 0.4 | -10.9* |
| 1990-2006 | -25.7* | 4.3 | -7.3* |
| 1990-2007 | -18.9* | 7.5 | -4.8* |
| 1990-2008 | -18.0* | 19.6 | -4.9* |
| 1990-2009 | -14.9* | 55.6 | -2.8* |
| 1990-2010 | -13.5* | 55.7 | -1.3 |
| 1990-2011 | -13.1* | 79.1 | -2.1* |
| 1990-2012 | -12.0* | 78.2 | -1.1 |
| 1990-2013 | -10.9* | 54.0 | 1.3 |
| 1990-2014 | -10.0* | 49.2 | 0.3 |
| 1990-2015 | -9.1* | 42.6 | 1.2 |
| 1990-2016 | -7.1* | 35.7 | 2.1 |

Columns 3-5 reports the values of the statistic proposed by Phillips and Sul (2007) to test for convergence for the female HIV prevalence rates (FPHIV), male HIV prevalence rates (MPHIV) and the total HIV incidence rates (TIHIV) of the Sub-Saharan countries included in Table 1.1.

Column 1 reports the different samples employed.

The PS statistic is distributed as a simple one-sided t-test with a critical value of −1.65. See Phillips and Sul (2007) for further details.

* indicates the rejection of the null hypothesis of convergence for the 5% significance level.

Table 1.3. Average values of the most important explanatory variables

| | Club1 | Club2 |
|---|---|---|
| Female primary persistence | 59.7 | 61.9 |
| Male primary persistence | 57.1 | 63.5 |
| Urban population | 31.6 | 36.4 |
| HDI | 0.48 | 0.42 |
| Per capita GDP | 3,794 | 2,383 |
| Female youth unemployment rate | 28.9 | 14.3 |
| Male youth unemployment rate | 26.4 | 11.3 |
| Female use of condom | 30.5 | 13.9 |

The countries included in clubs 1 and 2 have been determined in Section 2, although we have had to remove some of them due to the lack of data. The list of countries removed is in Table A1.

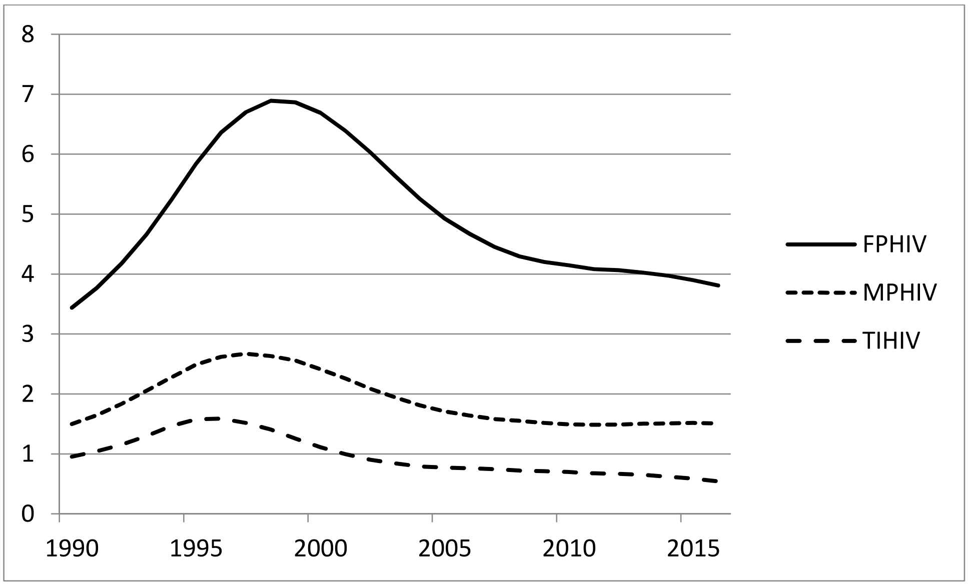
Table 1.4. Estimation of the probit model

| Variable | Estimated value (t-ratio) |
|---|---|
| Female Condom Use | 0.08 (2.5) |
| Female Primary Persistence | -0.06 (-3.2) |
| Urban Population | -0.05 (-2.2) |
| Youth Male Unemployment Rate | 0.06 (3.1) |
| Intercept | 2.35 (2.3) |
| Pseudo McFadden $R^2$ | 0.54 |
| % cases correctly classified (cutoff=0.5) | 92% |
| % cases correctly classified (cutoff=0.53) | 94% |

The dependent variable is constructed by assigning the value 0 to countries in club 2 and the value 1 to hose in club 1

The corresponding robust t-ratios are in parenthesis. All of them reject the Ho: $\beta_i$=0 (i=1-5) null hypothesis when a standard 5% significance level is employed.

Figure 1.1. σ-convergence




This figure represents the evolution of the standard deviation of the HIV female prevalence (FPHIV), male prevalence (MPHIV) and total incidence (TIHIV) rates.

Figure 1.2. Estimated FPHIV convergence clubs

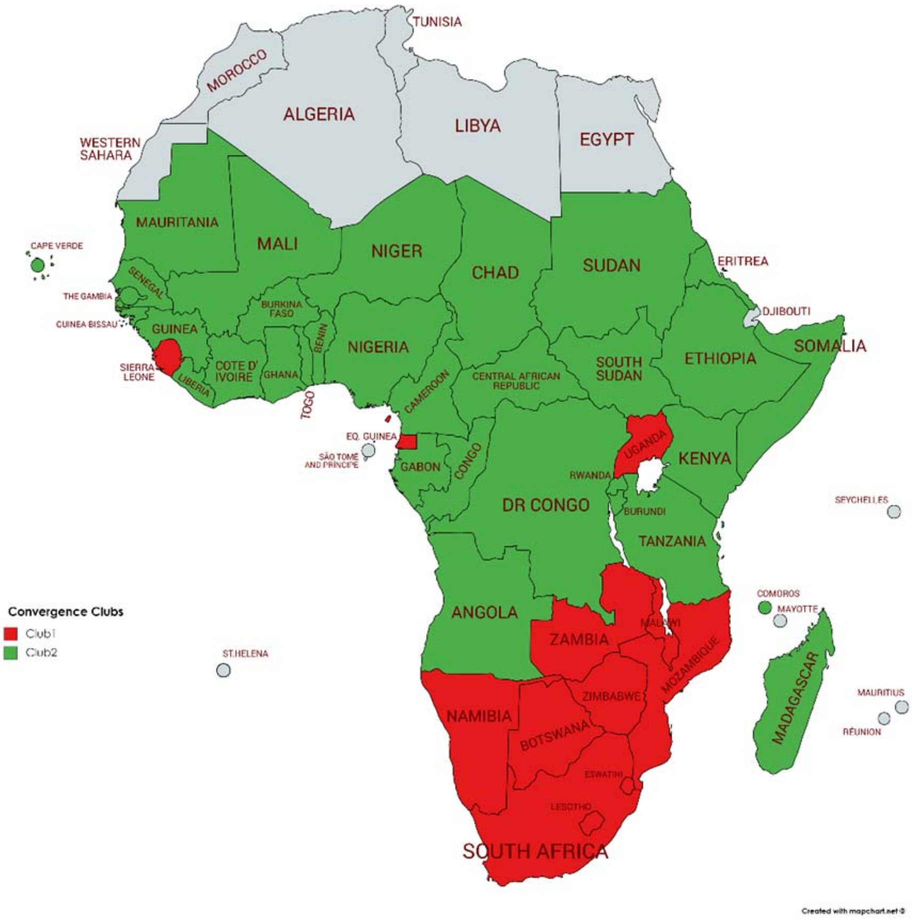
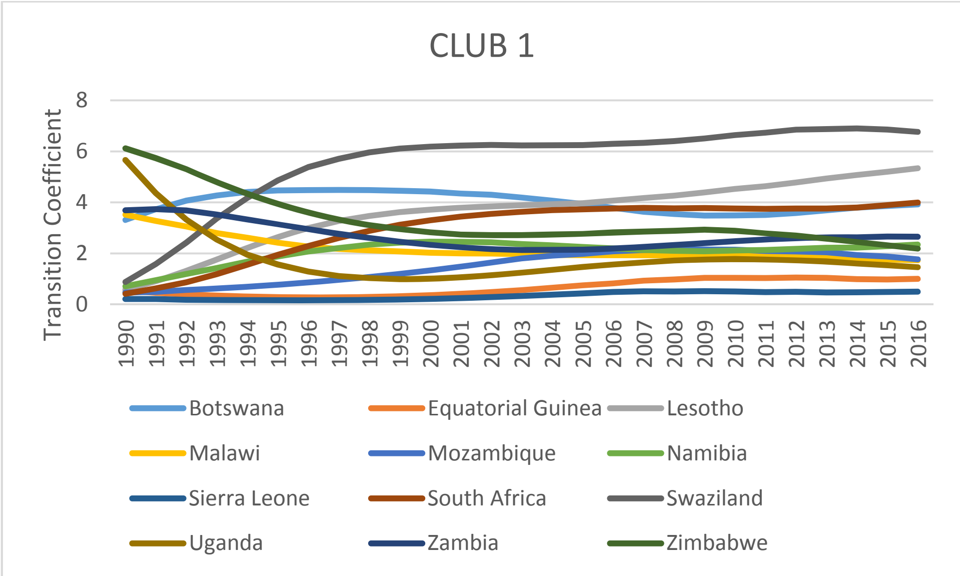
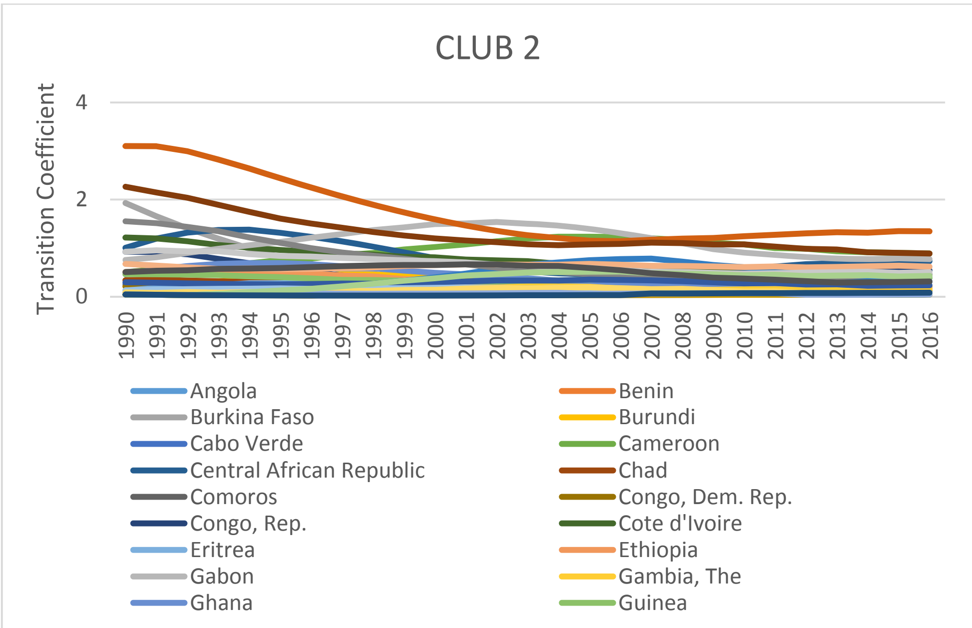


Figure 1.3. Transition coefficient (hit) evolution of FPHIV rates.

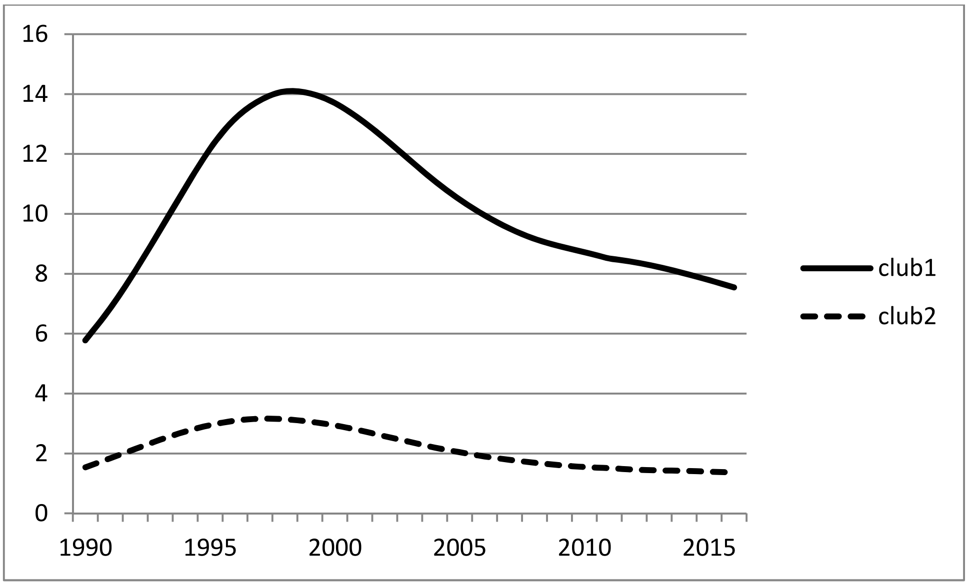



Transition coefficient ($h_{it}$) of FPHIV rates of the countries included in clubs 1.




Transition coefficient ($h_{it}$) of FPHIV rates of the countries included in clubs 2.

Figure 1.4. Average value of the estimated clubs




This figure presents the average values of FPHIV rates of the countries included in clubs 1 and 2.

Figure 1.5. Evolution of the probability of a country being included in club 1 as a function of the female persistence to the last year of primary.

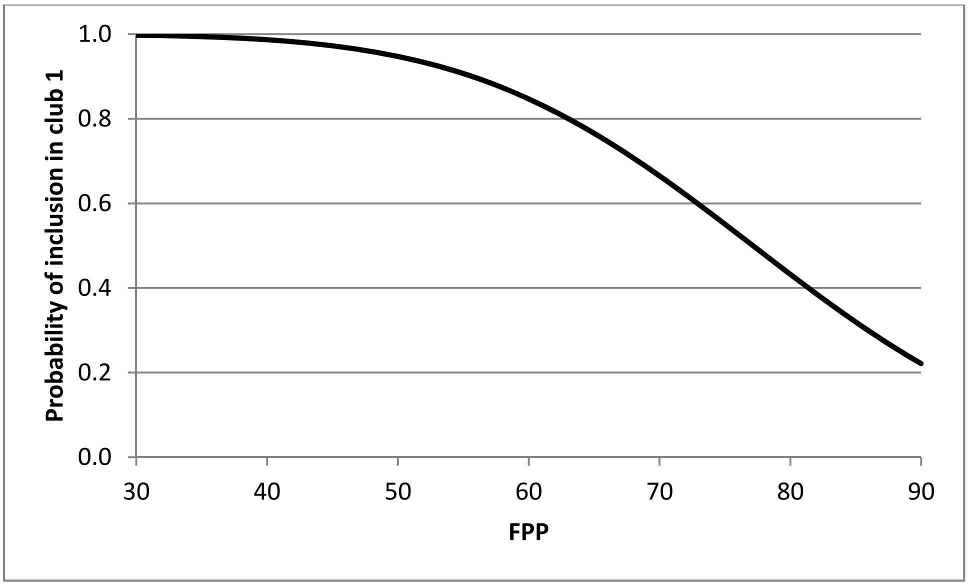
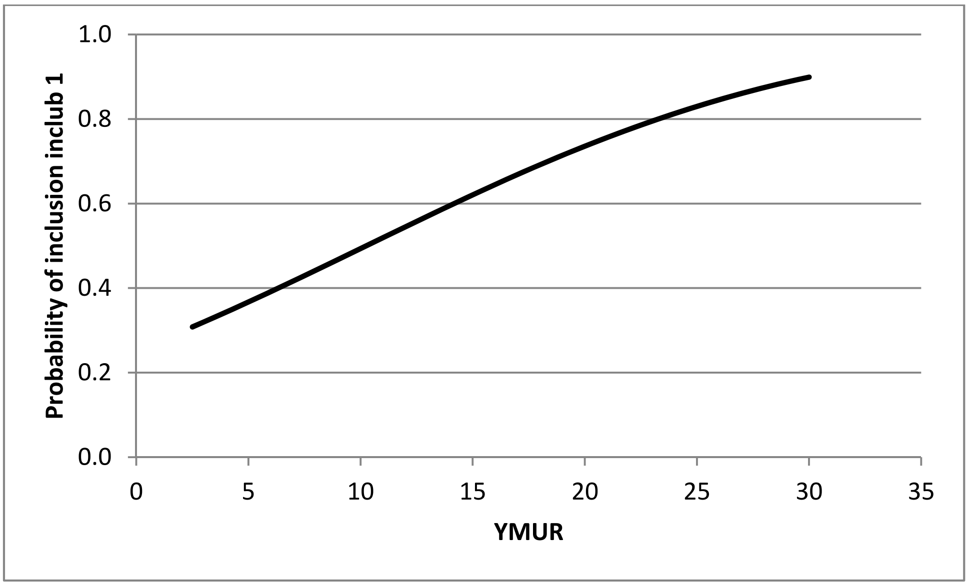


Figure 1.6. Evolution of the probability of a country being included in club as a function of the male youth unemployment rate.



# Chapter 2.

# Do international child mortality rates converge?

## 2.1. Introduction

The aim of policy-makers is to introduce appropriate measures to help the people of a country improve their living standard. To achieve this, the governments have to take decisions in several areas, like general economy, education and health. As a consequence, we can assess the status of a particular country by looking at the development of some socioeconomic indicators. This paper focuses on one of the most frequently used indicators to determine the health status of a particular society, namely, the evolution of child mortality. Its importance can be appreciated if we bear in mind that the United Nations considered the reduction of child mortality as one of the Millennium Development Goals (MDGs). This very ambitious plan ended in 2015 and was substituted by the Sustainable Development Goals (SDGs), officially known as Transforming Our World: the 2030 Agenda for Sustainable Development. This new agenda includes, in its Goal 3, called Good Health and Well-being, point 3.2, the reduction of child mortality rates.[2]

[2] Goal 3.2: "By 2030, end preventable deaths of newborns and children under 5 years of age, with all countries aiming to reduce neonatal mortality to at least as low as 12 per 1,000 live births and under-5 mortality to at least as low as 25 per 1,000 live births".

Following the data published in the United Nations webpage,[3] the reduction of child mortality has been remarkable, as can be observed in Figure 2.1. These positive results, however, cannot conceal the fact that 5.9 million children under the age of 5 died in 2015 and that, consequently, additional international effort is necessary. Moreover, it is not clear whether this huge improvement has been achieved similarly in all the countries or whether, by contrast, the effort has not been homogeneous and that, consequently, several patterns of behavior can be identified.

We should note that the evolution of child mortality rates has been previously analyzed in the literature, paying some attention to the possible presence of convergence. However, the conclusions are far from clear. Some papers, such as Wilson (2001) and Moser et al. (2005), conclude in favor of convergence in child mortality during 1950-2000, especially since 1980. Wilson (2001) states that the demographic differences are diminishing but that the economic distance between the countries remains the same. A very similar conclusion is drawn in Wilson (2011), where this author suggests the presence of different convergence clubs. Bremberg (2017) finds convergence between the OECD countries for 1990-2010. By contrast, McMichael et al. (2004) find a quite heterogeneous panorama where countries with rapid health gains co-exist with others that exhibit slower gains or even losses. Finally, Gil-Alana et al. (2017) study the time properties of international infant mortality rates. Using information about the number of infants who died before reaching the age of 1, they calculated per 1,000 live births in a given year for 37 countries and observed that the persistence of these rates varies substantially across countries. Consequently, under these circumstances, it is quite difficult to assume international convergence in infant mortality rates.

[3] http://www.un.org/millenniumgoals/childhealth.shtml

We should note that these previous results are mostly based on the use of the standard σ-convergence analysis, which offers evidence in favor of convergence whenever the cross-sectional dispersion of the data declines. However, none of these papers directly test for the null hypothesis of convergence. Thus, it seems appropriate to employ more advanced techniques, such as the methodology recently proposed in Phillips and Sul (2007, 2009) to directly test this hypothesis and determine whether international child mortality rates converge. Moreover, this method allows us to cluster the countries if the null hypothesis of convergence is rejected. Then, we can determine the existence of converge clubs, which means the presence of different patterns of behavior in international child mortality rates, and compare them.

Following this new approach we are in a position to shed light on the next issues:

- Are the MDGs / SDGs met homogenously in terms of reducing child mortality?
- Which countries have problems to achieve the MDGs / SDGs with respect to the objective established for the IMR?
- What are the socioeconomic determinants that limit a country in the attempt to reduce its IMR?

The rest of the paper is organized as follows. Section 2.2 describes the database employed. Section 2.3 tests for the convergence null hypothesis for child mortality rates. Section 2.4 determines the key variables that drive the creation of the clubs. The paper ends with a discussion of the results we have obtained.

## 2.2. Database description

In order to analyze the convergence of international child mortality rates, we take two different measurements, namely, the infant mortality rate (IMR) and the under-5 mortality rate (U5MR). The U5MR is defined as the number of deaths of children under 5 per 1000 live births, while the IMR refers to the children under one year of age. The data have been taken from the World Development Indicators of the World Bank. We have selected the 154 countries[4] that present a complete set of data for both measures for 1975-2015. Table 2.1 reflects these countries, their acronyms and some descriptive statistics for the two measures. These descriptive values allow us to observe that all the countries reveal negative growth rates in both mortality measures. Afghanistan, Guinea and Malawi present the highest rates of infant mortality in 1975, with 184.0, 184.5 and 186.6 deaths per 1000 live births. Similarly, Mali, Niger and Guinea show the highest rates of U5MR in 1975, with 368, 337 and 310 deaths per 1,000 live births. By contrast, Sweden, Finland and Japan present the lowest IMR in 1975, with 8.7, 10 and 10 deaths, while Sweden, Finland and Denmark have the lowest U5MR, with 10.3, 12.2 and 12.3 deaths, respectively.

The situation varies somewhat in 2015. The countries with the highest IMR are Central African Republic, Sierra Leone and Chad, with 91.5, 87 and 85 deaths per 1,000 live births, while the lowest rates are found in Luxembourg, Iceland and Finland, all of them well below 2 deaths per 1,000 live births. Similarly, Chad, Central African Republic and Sierra Leone are ranked in the worst positions for the U5MR, with 139, 130 and 120 deaths, while the top positions are occupied by Luxembourg, Iceland and Finland, with 1.9, 2.0 and 2.3 deaths per 1,000 live births. If we analyze the growth rates, we can observe that Maldives, Portugal and Oman present the highest reduction (in percentage points) in both IMR and U5MR. By contrast, Dominica and Central African Republic have an average growth rate of around a meager -1.0% for IMR, while Central African Republic and Zimbabwe present similar figures for the U5MR.

[4] The sample includes 47 African, 38 Asiatic, 32 American, 27 European and 10 Oceanic countries.

This last result is important, given that countries with high levels of child mortality at the beginning of the sample do not show larger reduction rates (in absolute values) than countries with low rates at the beginning of the sample. Consequently, convergence is quite dubitable, at least in terms of the standard β-convergence and σ-convergence concepts used in the seminal papers of Barro and Sala-i-Martí (1991, 1992). This is verified in Figure 2.2., which reflects the evolution of the variation coefficient across the sample of the two child mortality rates. As we can see, none of these coefficients goes towards 0, which again casts doubts about the presence of convergence. We can also see that the dispersion of the IMR is slightly lower than that of the U5MR and that both coefficients present similar growth rates until 1995. However, the evolution after this date is a bit different: the average growth rate of the IMR for 1995-2015 is 0.3, while it is 0.0 for the U5MR. Moreover, the variation coefficient of both rates is slightly negative (-0.1) during the new century. So, this result can be interpreted in a positive way, stating that it might imply a general deceleration of the highest mortality rates and their approximation towards the lowest rates. If this were true, convergence in child mortality rates could be achieved in the future. In any event, disparities still persist nowadays and it seems to be appropriate to analyze them to determine the presence of convergence.

### 2.3. Testing for convergence in child mortality

In order to test for the null hypothesis of convergence, we follow the recent papers of Phillips and Sul (2007, 2009) (PS, hereafter) where they develop a framework that allows us to test for the convergence hypothesis and then, if this hypothesis is rejected, to estimate the convergence clubs that group the child mortality rate behavior. The PS methodology is closely related to the standard σ-convergence analysis, a concept introduced by Barro and Sala-i-Martin (1990) and Barro et al. (1991), in that it tests for the decline of the variable of interest over time in the cross-sectional dispersion. However, it clearly outperforms the classical convergence analysis, namely, the abovementioned σ-convergence concept and the β-convergence concept, which was first introduced by Baumol (1986). The PS methodology is based on a general nonlinear time-varying factor model, which admits the presence of transitional heterogeneity. Additionally, it is flexible with respect to the time properties of the variables under analysis because it does not impose any particular assumption about them. Finally, it is clearly free of the criticism received by the β-convergence analysis in De Long (1988) and Quah (1993). As a consequence, the use of this methodology has recently become very popular in convergence analysis.

Following these authors, let us consider that $X_{it}$ represents either of the two child mortality rates considered in this paper, U5MR and IMR, with i=1, 2, …, 154 (the 154 previously selected countries) and t= 1975, …, 2015. This variable can be decomposed as $X_{it} = \delta_{it}\ \mu_t$, where $\mu_t$ and $\delta_{it}$ are the common and the idiosyncratic components, respectively. PS suggest testing for convergence by analyzing whether $\delta_{it}$ converges towards δ. To do so, they first define the relative transition component or transition coefficient:

$$h_{it} = \frac{X_{it}}{N^{-1}\ \sum_{i=1}^{N} X_{it}} = \frac{\delta_{it}}{N^{-1}\ \sum_{i=1}^{N} \delta_{it}} \tag{2.1}$$

In the presence of convergence, $h_{it}$ should converge towards unity, while its cross-sectional variance, $H_{it}$, is defined as follows:

$$H_{it} = N^{-1}\ \sum_{i=1}^{N} (h_{it} - 1)^2 \tag{2.2}$$

and should go to 0 when T goes towards infinity. Then, PS test for convergence by estimating the following equation:

$$log\frac{H_1}{H_t} - 2log[\log(t)] = \alpha + \beta\log(t) + u_t,\ t = T_o, \dots, T \qquad (2.3)$$

with $T_o = [rT]$, and r=0.3, as suggested by PS. Equation (2.3) is commonly known as the log-t regression. The null hypothesis of convergence is rejected whenever parameter β is lower than 0. PS suggest estimating model (2.3) by methods which correct for the presence of autocorrelation and heteroscedasticity and, later, employ the t-statistic to test the null hypothesis β=0. The use of these robust methods ensures that this t-ratio converges towards a standard N(0,1) distribution and, therefore, we will reject the null hypothesis of convergence whenever this t-statistic takes values lower than -1.65.

If we reject convergence, PS propose the following robust clustering algorithm for identifying clubs in a panel of countries:

i. Order the N countries according to their final values

ii. Starting from the highest-order countries, add adjacent countries from our ordered list and estimate model (2.3). Then, select the core group by maximizing the value of the convergence t-statistic, subject to the restriction that it is greater than -1.65.

iii. Continue adding one country at a time of the remaining countries to the core group, and re-estimate model (2.3) for each formation. Use the sign criterion (t-statistic >0) to decide whether a country should join the core group.

iv. For the remaining countries, repeat steps (ii)–(iii) iteratively and stop when clubs can no longer be formed. If the last group does not have a convergence pattern, conclude that its members diverge.

PS recommend performing club merging tests after running the algorithm using equation (2.3) in order to avoid the over-estimation of the number of clubs.

Finally, we have followed the suggestion of PS and we have extracted the trend components of the series by filtering them using the Hodrick and Prescott (1997) filter, applying the standard value $\lambda$=400.

The results of the convergence test are presented in Table 2.2. We can observe that we can reject the null hypothesis of convergence when the total sample is considered for the two child mortality rates. If we only take into account the pre Great Recession sample, we can again reject this null hypothesis. The estimated value of the $\beta$ parameter barely varies, which leads us to some intuition about the inexistence of significant changes in the convergence results related to the Great Recession. Therefore, as could be expected, the IMR do not converge towards a single pattern of behavior, but the existence of different convergence clubs is possible. In order to define these clubs, we can employ the PS algorithm. Tables 2.3 and 2.4 reflect the results for the two child mortality rates when the total sample is considered.

The analysis of the composition of the convergence clubs reveals some interesting insights. With respect to the IMR, in Table 2.3., we can observe that the African countries are mostly in clubs 1 and 2, the ones with the highest values. The exceptions are Algeria, Cabo Verde, Morocco and Seychelles, which are in club 3; Libya, Mauritius and Tunisia, in club 4. The Oceanic countries are mostly in clubs 2 and 3. Papua is in club 1, while New Zealand and Australia belong to clubs 5 and 6, respectively. The American and Asiatic countries show a similar distribution. They are mostly in clubs 3-5, which are closer to the mean of the distribution. Some of them, however, such as Afghanistan and Pakistan, and the American island of Dominica, are in club 1. We should also note that Cuba is assigned to club 6, together with Israel and Korea, and that Japan and Singapore are in club 8.

Finally, the European countries are mostly assigned to clubs 6-9, but some exceptions should be noted. For instance, Bulgaria, Georgia, Moldova, Romania and Ukraine are in club 4, while Hungary and Malta are in club 5. By contrast, Iceland and Luxembourg compose the club 9. Finally, we should note that Italy and Portugal are in club 7. Figure 2.3 plots the transition coefficient ($h_{it}$) for each club of IMR convergence.

The results for the U5MR are presented in Table 2.4 and are similar to the ones obtained for IMR. However, Club 1 of IMR is now split into two different clubs because of the very high rates of mortality of Central African Republic, Chad, Lesotho, Mali, Niger, Sierra Leone, Swaziland and Zimbabwe. Another difference is in the behavior of Luxembourg and Italy whose rates diverge and who are not included in any of the estimated convergence clubs. Figure 2.4 reproduce the transition coefficient ($h_{it}$) for each club of U5MR convergence.

We can also study the evolution of the estimated clubs by constructing indexes which reflect the average value of the child mortality rates of the countries in the different clubs. These indexes are reported in Figures 2.5.a and 2.5.b, where we can observe the efforts of each country to improve the health of its inhabitants. For the IMR, we can see that the distance between the estimated clubs has been reduced from 119 at the beginning of the sample to just 60 in 2015. Nevertheless, we must admit that this distance is still too large and that the differences in IMR are important. If we focus on the distance between clubs 3-5 and clubs 6-9, we observe that clubs 3 and 4 have reduced their mortality rate by more than 50, while club 5 has done so only by 33.

The results for the U5MR are somewhat similar because all the convergence clubs show a clear decreasing trend and the distance between the clubs has been reduced from 222 in 1975 to 101 in 2015. The composition of the clubs is also quite similar to the one observed for the IMR. The African countries are in the first club, while the European countries belong to clubs 5-10.

### 2.4. Which are the forces that drive the club creation?

The previous section has shown that the IMR and U5MR do not follow a single pattern of behavior across the whole world. On the contrary, we observe multiple patterns. Consequently, we should now try to identify the sources of these differences. We have to bear in mind that some authors as Scell et al. (2007) and Muldoon et al. (2001) has suggested several factors to explain these differences in child mortality rates. Some economic variables are commonly used to explain the differences in child mortality including the per capita Gross Domestic Product (GDP), as we can see in the papers of Gbesemete and Jonsson (1993), Alves and Belluzzo (2004), Hakobyan et al. (2006), Renton et al. (2012) and Subramaniam et al. (2018). Ko et al. (2014) alternatively consider the employment rate.

From a different perspective, some authors use educational levels as another important factor to understand the evolution of child mortality. In this regard, we can cite the papers of Mondal et al. (2009) and Defo (1996). Others also find an interesting relationship between child mortality and female fertility, some recent analyses, such as Hondroyiannis and Papapetrou (2002) for Finland and Narayan and Peng (2007) for Japan, finding a negative relationship between these two variables.

Additional factors that should be taken into account are those offered by the demographic perspective. There is a large body of literature which associates health with place of residence. For instance, Jankowska et al. (2014), Fink et al. (2014) and Quansah et al. (2016) have recently found evidence of the comparative advantages of urban residence over rural residence, mainly due to the improvement in living standards offered by urban zones, like access to basic sanitation services, education and housing. However, we should note that the influence of urban population on IMR seems to have decreased in recent years, if we consider the results of Kimani-Murage et al. (2014). Similarly, Hathi et al. (2014) find a relationship between population density and infant mortality.

Finally, we should consider that some authors have found a close relationship between child mortality and access to healthcare. This result is supported by Hanmer et al. (2003), amongst others.

In order to reflect all these potential explanatory factors, the variables that we have selected are presented in Table 2.5. The socioeconomic variables selected are per capita GDP, unemployment rate, and the female literacy rate; health factors are measured by access to basic sanitary services, percentage of births attended by skilled staff and health expenditure; the demographic factors are fertility rate, percentage of urban population and population density; political conditions are number of peacekeepers in the country, number of deaths in battle and intentional homicides. Additionally, we have considered some geographical dummy variables $D_{ij}$, i={Africa, America, Asia, Europe, Oceania} which take the value 1 if the t-th country belongs to the i-th continent and 0 otherwise. Similarly, we have created the dummy variable $D_{OECDj}$ that takes the value 1 if the country belongs to OECD at the end of the sample and 0 otherwise.

All the data have been taken from the World Development Indicators database, further details about the sources of the data are presented in Table B1. The sample covers 1975-2015 and we have employed the latest available data for most of the explanatory variables. Exceptions are the number of battle-related deaths, the number of peacekeepers and the number of intentional homicides. We have employed the sample average for them. Having selected the set of possible explanatory variables, we subsequently estimate an ordered probit model in which the dependent variable is constructed in such a manner that it takes the value i if the country is assigned to the i-th club. The final specification is selected using a general-to-particular strategy where the variables are iteratively removed when their estimated coefficient is not statistically different to 0. The results are presented in Table 2.6.

The estimated models for the two child mortality rates are quite similar so we will focus on the analysis of the model for IMR for the sake of brevity, given that the comments for the U5MR estimated model would be extremely similar. The results in Table 2.6 show the importance of the economic variables in explaining the level of child mortality. We should note that increases in the per capita GDP and reductions in the total unemployment rate reduces the probability of a country being assigned to club 1. We also considered the female and male unemployment rates, but these variables do not improved the estimation. Furthermore, increases in access to basic sanitary services and public health expenditure also reduces the probability of being included in club 1. We should note that our results confirm the inverse relationship between fertility and child mortality. Finally, we can observe the importance of the demographic factors. The estimated coefficients of both population density and percentage of urban population are highly significant, proving that the concentration of population in cities helps to reduce the child mortality rates.

In order to evaluate the importance of the explanatory variables, we have also included the estimated elasticities. We can observe that access to basic sanitation services takes the highest value. This result is no surprise, given that the average value of this variable for the countries in club 1 is only 32.26, far from those of clubs 5-9 (98.4). Therefore, this variable is extremely useful to determine into which club the countries are classified and increases in this variable more than proportionally increment the probability of a country moving to the clubs with the lowest child mortality rates. Similarly, the elasticities of the economic variables (per capita GDP, total unemployment rate and public health expenditure) are greater than 2 (in absolute terms) so we can conclude that the reduction of child mortality rates should be based on improvements of the values of these variables.

Finally, we analyze how the estimated probabilities change when the explanatory variables vary. We should note that, if we know the variation in the estimated probabilities, then we can analyze the evolution of the predicted child mortality rates by simply multiplying the estimated probabilities by the average values of the IMR of each club. To do this, we need information about the behavior of the explanatory variables. Let us focus on the behavior of the countries of club 1, given that they present higher child mortality rates than those pursued by the SDGs and it is questionable whether these countries will be able to achieve these goals. We should begin by taking the average values of the explanatory variables of the countries in club 1 and allowing them to vary under the following assumptions:[5] the per capita GDP grows at rate (1+g); public health expenditure, access to basic sanitation services, the total unemployment rate and the fertility rate depend on the per capita GDP; and the population density and the percentage of urban population grows at 1%. We obtain the estimated probabilities

[5] See more details in Appendix B2.

for the new values of the explanatory variables. Figure 2.6 reflects the variation of these estimated probabilities when g takes values in the (0, 5%) interval. As can be seen, the probability of moving from club 1 to the rest of the clubs clearly depends on g: the higher the value of this parameter, the more probable the change from club 1 to the other clubs. An increase of 60% in the per capita GDP is necessary for the estimated probability of being in club 2 to exceed the probability of being in club 1. For this to occur, the per capita GDP must be greater than 1,901$. Similarly, for the estimated probability of being in club 3 to exceed 50%, the per capita GDP must be greater than 5,183$, which implies a growth rate of more than 340%. Finally, even if we multiply the per capita GDP by 5, then the probabilities of being included in clubs 2, 3 and 4 are 0.03, 0.54 and 0.42, respectively.

These estimated probabilities can be translated into the corresponding mortality rates. If we multiply the estimated probabilities by the average values of the IMR of each club, then the estimated infant mortality rate is 16.1, which is a remarkable reduction from the average IMR value of the countries in club 1 in 2015 (61).

The previous analysis shows how the better the socioeconomic conditions, the lower the IMR. However, we should note that the predicted IMR is far from the value pursued in the SDGs, namely, an IMR lower than 12 deaths per 1,000 live births. Thus, we look for the value of the parameter g that could lower the IMR to less than 12 in 2030. Under the assumptions previously considered, only a value of g greater than 14.95% could lower the IMR of an average-type club 1 country to below 12 in 2030. In this case, the values of the explanatory variables would be $9,632 for per capita GDP, 5.5 percent for total unemployment rate, 11.6 percent for public health expenditure, as a percentage of the GDP, 87 for population density, 44 percent for urban population, 1.8 for total fertility rate and 88.1 the percentage of population with access to basic sanitation services. If we consider longer horizons, we can see that the IMR target would be possible for 2050 with a yearly growth rate of higher than 6%. Per capita GDP growth rates of this size are far from the historical growth rates of these countries and, consequently, it is almost impossible to fulfil the reduction of the child mortality rates stated in the SDGs.

### 2.5. Conclusions

This paper analyses the convergence of worldwide child mortality rates. We consider two different measurements, the infant mortality rate and the under-5 mortality rate. Our results lead us to reject the existence of a unique pattern of behavior in both rates. The evolution of these rates for the countries considered can, in fact, be grouped into up to 10 different convergence clubs. The African and Asiatic countries are mostly included in the clubs with the highest mortality rates, while the European countries belong to the clubs with the lowest rates, although the rates of Cuba, Korea, Japan and Singapore are similar to those of the European countries.

When analyzing why a particular country is included in a club, the results of the estimation of an ordered probit model show that this classification depends on population density, fertility rate, per capita GDP, access to basic sanitation services, percentage of urban population and total unemployment rate. Our results confirm the existence of an inverse relationship between fertility and child mortality rate.

Finally, our conclusions offer interesting results for policy-makers. We have seen that the creation of a club depends on the economic situation and, especially, on access to basic sanitation services. Furthermore, we have analyzed how variations in the explanatory variables change the estimated probabilities of being included in the different clubs and how these variations can help us to predict the values of the child mortality rates. The results of this paper lead us to conclude that it is unlikely that the

SDGs will be achieved in 2030. The reduction of the IMR to below 12 deaths per 1,000 live births would require unlikely increments of the per capita GDP. Even if the fulfilment of the goals were postponed to 2050, the economies would have to grow at very improbable rates.

## Table 2.1. List of countries and descriptive analysis

| Code | Country Name | Acronym | IMR75 | IMR15 | $g_{IMR}$ | U5MR75 | U5MR15 | $g_{U5MR}$ |
|---|---|---|---|---|---|---|---|---|
| 1 | Afghanistan | AFG | 186.6 | 66.3 | -2.5% | 279.4 | 91.1 | -2.7% |
| 2 | Algeria | DZA | 128.7 | 21.9 | -4.2% | 205 | 25.5 | -5.0% |
| 3 | Argentina | ARG | 53.2 | 11.1 | -3.8% | 63.8 | 12.5 | -3.9% |
| 4 | Australia | AUS | 14.4 | 3 | -3.8% | 17.3 | 3.8 | -3.6% |
| 5 | Austria | AUT | 19.8 | 2.9 | -4.6% | 23.1 | 3.5 | -4.5% |
| 6 | Bahamas, The | BHS | 27.2 | 9.9 | -2.4% | 32.2 | 12.1 | -2.4% |
| 7 | Bahrain | BHR | 37.7 | 5.3 | -4.7% | 47.4 | 6.2 | -4.8% |
| 8 | Bangladesh | BGD | 145.2 | 30.7 | -3.7% | 217.1 | 37.6 | -4.2% |
| 9 | Barbados | BRB | 31.8 | 12 | -2.3% | 36.7 | 13 | -2.5% |
| 10 | Belgium | BEL | 15.8 | 3.3 | -3.7% | 18.6 | 4.1 | -3.6% |
| 11 | Belize | BLZ | 59.2 | 14.2 | -3.4% | 79.9 | 16.5 | -3.8% |
| 12 | Benin | BEN | 143.5 | 64.2 | -1.9% | 242.7 | 99.5 | -2.2% |
| 13 | Bhutan | BTN | 157.1 | 27.2 | -4.2% | 235.5 | 32.9 | -4.7% |
| 14 | Bolivia | BOL | 127.9 | 30.6 | -3.4% | 200.1 | 38.4 | -3.9% |
| 15 | Botswana | BWA | 69.5 | 34.8 | -1.7% | 96.1 | 43.6 | -1.9% |
| 16 | Brazil | BRA | 89.8 | 14.6 | -4.3% | 115.3 | 16.4 | -4.6% |
| 17 | Bulgaria | BGR | 29.8 | 9.3 | -2.8% | 34.3 | 10.4 | -2.9% |
| 18 | Burkina Faso | BFA | 136.6 | 60.9 | -2.0% | 288.5 | 88.6 | -2.8% |
| 19 | Burundi | BDI | 145.7 | 54.1 | -2.4% | 246.5 | 81.7 | -2.7% |
| 20 | Cabo Verde | CPV | 78 | 20.7 | -3.2% | 109.4 | 24.5 | -3.6% |
| 21 | Cambodia | KHM | 178.2 | 24.6 | -4.7% | 310 | 28.7 | -5.6% |
| 22 | Cameroon | CMR | 114.3 | 57.1 | -1.7% | 191.1 | 87.9 | -1.9% |
| 23 | Canada | CAN | 14.2 | 4.3 | -2.9% | 17 | 4.9 | -3.0% |
| 24 | Ctral. African Rep. | CAF | 126 | 91.5 | -0.8% | 199.3 | 130.1 | -1.0% |
| 25 | Chad | TCD | 130 | 85 | -1.0% | 248.1 | 138.7 | -1.4% |
| 26 | Chile | CHL | 55.3 | 7 | -4.9% | 64.7 | 8.1 | -4.9% |
| 27 | China | CHN | 62.6 | 9.2 | -4.6% | 85.2 | 10.7 | -4.9% |
| 28 | Colombia | COL | 58.4 | 13.6 | -3.5% | 78.7 | 15.9 | -3.8% |
| 29 | Comoros | COM | 136.2 | 55.1 | -2.2% | 202.7 | 73.5 | -2.4% |
| 30 | Congo, Dem. Rep. | COD | 140.7 | 74.5 | -1.5% | 230.3 | 98.3 | -2.1% |
| 31 | Congo, Rep. | COG | 79.2 | 33.2 | -2.1% | 126.3 | 45 | -2.5% |
| 32 | Costa Rica | CRI | 38 | 8.5 | -3.6% | 46.3 | 9.7 | -3.7% |
| 33 | Cote d'Ivoire | CIV | 135.1 | 66.6 | -1.7% | 201.1 | 92.6 | -1.9% |
| 34 | Cuba | CUB | 24.7 | 4 | -4.3% | 29.4 | 5.5 | -4.0% |
| 35 | Denmark | DNK | 10.3 | 2.9 | -3.0% | 12.3 | 3.5 | -3.0% |
| 36 | Dominica | DMA | 25.5 | 19.6 | -0.6% | 34.5 | 21.2 | -1.2% |
| 37 | Dominican Republic | DOM | 74.9 | 25.7 | -2.6% | 104.5 | 30.9 | -2.9% |
| 38 | Ecuador | ECU | 82.5 | 18.4 | -3.6% | 116.6 | 21.6 | -4.0% |
| 39 | Egypt, Arab Rep. | EGY | 143.7 | 20.3 | -4.7% | 214.7 | 24 | -5.2% |
| 40 | El Salvador | SLV | 95.6 | 14.4 | -4.5% | 137.4 | 16.8 | -5.0% |
| 41 | Eritrea | ERI | 123.9 | 34.1 | -3.1% | 208.8 | 46.5 | -3.6% |

| Code | Country Name | Acronym | IMR75 | IMR15 | $g_{IMR}$ | U5MR75 | U5MR15 | $g_{U5MR}$ |
|---|---|---|---|---|---|---|---|---|
| 42 | Ethiopia | ETH | 142.7 | 41.4 | -3.0% | 241.5 | 59.2 | -3.4% |
| 43 | Fiji | FJI | 41.4 | 19.1 | -1.9% | 52.8 | 22.4 | -2.1% |
| 44 | Finland | FIN | 10 | 1.9 | -4.0% | 12.2 | 2.3 | -4.0% |
| 45 | France | FRA | 12.4 | 3.5 | -3.0% | 15 | 4.3 | -3.0% |
| 46 | Gambia, The | GMB | 114.4 | 47.9 | -2.1% | 270.4 | 68.9 | -3.3% |
| 47 | Georgia | GEO | 56.6 | 10.6 | -4.0% | 68.6 | 11.9 | -4.2% |
| 48 | Germany | DEU | 17.7 | 3.1 | -4.2% | 20.8 | 3.7 | -4.1% |
| 49 | Ghana | GHA | 111.5 | 42.8 | -2.3% | 185.9 | 61.6 | -2.7% |
| 50 | Greece | GRC | 28.5 | 3.6 | -4.9% | 31.4 | 4.6 | -4.6% |
| 51 | Guatemala | GTM | 102.9 | 24.3 | -3.5% | 149 | 29.1 | -3.9% |
| 52 | Guinea | GIN | 184.5 | 61 | -2.7% | 310.1 | 93.7 | -2.9% |
| 53 | Guyana | GUY | 54.6 | 32 | -1.3% | 72.7 | 39.4 | -1.5% |
| 54 | Haiti | HTI | 146 | 52.2 | -2.5% | 218.4 | 69 | -2.8% |
| 55 | Honduras | HND | 85 | 17.4 | -3.8% | 120.5 | 20.4 | -4.2% |
| 56 | Hungary | HUN | 35.8 | 5.3 | -4.6% | 39.2 | 5.9 | -4.5% |
| 57 | Iceland | ISL | 10.3 | 1.6 | -4.4% | 12.8 | 2 | -4.4% |
| 58 | India | IND | 130.3 | 37.9 | -3.0% | 193.4 | 47.7 | -3.4% |
| 59 | Indonesia | IDN | 97.8 | 22.8 | -3.5% | 140.8 | 27.2 | -3.9% |
| 60 | Iran, Islamic Rep. | IRN | 105.8 | 13.4 | -4.9% | 153.7 | 15.5 | -5.4% |
| 61 | Iraq | IRQ | 66.2 | 26.5 | -2.2% | 90.8 | 32 | -2.5% |
| 62 | Ireland | IRL | 17 | 3 | -4.1% | 20 | 3.6 | -4.1% |
| 63 | Israel | ISR | 22.4 | 3.2 | -4.6% | 26.2 | 4 | -4.5% |
| 64 | Italy | ITA | 20.5 | 2.9 | -4.7% | 23.1 | 3.5 | -4.5% |
| 65 | Jamaica | JAM | 37.8 | 13.5 | -2.5% | 47.6 | 15.7 | -2.7% |
| 66 | Japan | JPN | 10 | 2 | -3.8% | 13.3 | 2.7 | -3.8% |
| 67 | Jordan | JOR | 53.3 | 15.4 | -3.0% | 70.6 | 17.9 | -3.3% |
| 68 | Kazakhstan | KAZ | 63.4 | 12.6 | -3.9% | 77.8 | 14.1 | -4.1% |
| 69 | Kenya | KEN | 80.7 | 35.5 | -2.0% | 129 | 49.4 | -2.3% |
| 70 | Kiribati | KIR | 88.5 | 43.6 | -1.7% | 126 | 55.9 | -2.0% |
| 71 | Korea, Rep. | KOR | 22.8 | 2.9 | -4.9% | 27.1 | 3.4 | -4.9% |
| 72 | Kuwait | KWT | 40.5 | 7.3 | -4.1% | 51.4 | 8.6 | -4.3% |
| 73 | Kyrgyz Republic | KGZ | 87.9 | 19 | -3.7% | 112.6 | 21.3 | -4.0% |
| 74 | Lebanon | LBN | 45.4 | 7.1 | -4.4% | 58.6 | 8.3 | -4.7% |
| 75 | Lesotho | LSO | 114.6 | 69.2 | -1.2% | 150.8 | 90.2 | -1.2% |
| 76 | Liberia | LBR | 174.2 | 52.8 | -2.9% | 261.2 | 69.9 | -3.2% |
| 77 | Libya | LBY | 72.8 | 11.4 | -4.4% | 95.5 | 13.4 | -4.7% |
| 78 | Lithuania | LTU | 20.2 | 3.3 | -4.3% | 24.8 | 5.2 | -3.7% |
| 79 | Luxembourg | LUX | 14.2 | 1.5 | -5.3% | 16.7 | 1.9 | -5.2% |
| 80 | Madagascar | MDG | 99.5 | 35.9 | -2.5% | 163.5 | 49.6 | -2.9% |
| 81 | Malawi | MWI | 184 | 43.4 | -3.5% | 309.4 | 64 | -3.8% |
| 82 | Malaysia | MYS | 34.2 | 6 | -4.2% | 42.5 | 7 | -4.3% |
| 83 | Maldives | MDV | 137.2 | 7.4 | -6.9% | 204.4 | 8.6 | -7.4% |
| 84 | Mali | MLI | 180.9 | 74.5 | -2.1% | 368.3 | 114.7 | -2.8% |
| 85 | Malta | MLT | 19 | 5.1 | -3.2% | 21.4 | 6.4 | -2.9% |

| Code | Country Name | Acronym | IMR75 | IMR15 | $g_{IMR}$ | U5MR75 | U5MR15 | $g_{U5MR}$ |
|---|---|---|---|---|---|---|---|---|
| 86 | Marshall Islands | MHL | 57.5 | 29.6 | -1.6% | 77.3 | 36 | -1.8% |
| 87 | Mauritania | MRT | 105.1 | 65.1 | -1.2% | 185.5 | 84.7 | -1.9% |
| 88 | Mauritius | MUS | 49.6 | 11.8 | -3.4% | 65.2 | 13.5 | -3.8% |
| 89 | Mexico | MEX | 66.6 | 11.3 | -4.2% | 91.5 | 13.2 | -4.6% |
| 90 | Moldova | MDA | 44.9 | 13.6 | -2.9% | 57.9 | 15.8 | -3.1% |
| 91 | Morocco | MAR | 109.9 | 23.7 | -3.7% | 166.2 | 27.6 | -4.3% |
| 92 | Mozambique | MOZ | 178 | 56.7 | -2.8% | 266.7 | 78.5 | -2.9% |
| 93 | Myanmar | MMR | 109.1 | 39.5 | -2.4% | 159.1 | 50 | -2.8% |
| 94 | Namibia | NAM | 63.4 | 32.8 | -1.6% | 98.1 | 45.4 | -1.9% |
| 95 | Nepal | NPL | 159.9 | 29.4 | -4.0% | 239.8 | 35.8 | -4.5% |
| 96 | Netherlands | NLD | 10.5 | 3.2 | -2.9% | 13.1 | 3.8 | -3.0% |
| 97 | New Zealand | NZL | 14.9 | 4.7 | -2.8% | 18.3 | 5.7 | -2.8% |
| 98 | Nicaragua | NIC | 100.3 | 18.8 | -4.0% | 144.7 | 22.1 | -4.5% |
| 99 | Niger | NER | 141.4 | 57.1 | -2.2% | 337.1 | 95.5 | -3.0% |
| 100 | Nigeria | NGA | 144.4 | 69.4 | -1.8% | 244.3 | 108.8 | -2.0% |
| 101 | Norway | NOR | 10.5 | 2 | -4.0% | 12.9 | 2.6 | -3.8% |
| 102 | Oman | OMN | 114.2 | 9.9 | -5.8% | 167.4 | 11.6 | -6.3% |
| 103 | Pakistan | PAK | 129.4 | 65.8 | -1.6% | 172 | 81.1 | -1.8% |
| 104 | Panama | PAN | 41.8 | 14.6 | -2.5% | 53.3 | 17 | -2.7% |
| 105 | Papua New Guinea | PNG | 86.9 | 44.5 | -1.6% | 123.5 | 57.3 | -1.9% |
| 106 | Paraguay | PRY | 54.5 | 17.5 | -2.7% | 72.5 | 20.5 | -3.0% |
| 107 | Peru | PER | 91 | 13.1 | -4.6% | 140.5 | 16.9 | -5.0% |
| 108 | Philippines | PHL | 54.4 | 22.2 | -2.2% | 82.1 | 28 | -2.6% |
| 109 | Poland | POL | 24.5 | 4.5 | -4.0% | 27.7 | 5.2 | -4.0% |
| 110 | Portugal | PRT | 36 | 3 | -5.9% | 43.5 | 3.6 | -5.9% |
| 111 | Qatar | QAT | 37.7 | 6.8 | -4.1% | 47.5 | 8 | -4.3% |
| 112 | Romania | ROU | 42.8 | 9.7 | -3.6% | 50.4 | 11.1 | -3.6% |
| 113 | Russian Federation | RUS | 29.9 | 8.2 | -3.1% | 36.6 | 9.6 | -3.2% |
| 114 | Rwanda | RWA | 144.8 | 31.1 | -3.7% | 244.9 | 41.7 | -4.2% |
| 115 | Sao Tome and Principe | STP | 58.9 | 34.6 | -1.3% | 90.1 | 47.3 | -1.6% |
| 116 | Saudi Arabia | SAU | 93.1 | 12.5 | -4.8% | 133.4 | 14.5 | -5.3% |
| 117 | Senegal | SEN | 108 | 41.7 | -2.3% | 252.4 | 47.2 | -4.0% |
| 118 | Seychelles | SYC | 39.7 | 11.7 | -2.9% | 50.4 | 13.6 | -3.1% |
| 119 | Sierra Leone | SLE | 177.5 | 87.1 | -1.7% | 308.7 | 120.4 | -2.3% |
| 120 | Singapore | SGP | 14.3 | 2.1 | -4.6% | 17.6 | 2.7 | -4.5% |
| 121 | Solomon Islands | SLB | 55.1 | 23.6 | -2.0% | 73.5 | 28.1 | -2.3% |
| 122 | South Africa | ZAF | 86.6 | 33.6 | -2.3% | 126.6 | 40.5 | -2.7% |
| 123 | Spain | ESP | 20.4 | 3.5 | -4.2% | 23.5 | 4.1 | -4.2% |
| 124 | Sri Lanka | LKA | 48.7 | 8.4 | -4.2% | 63.6 | 9.8 | -4.5% |
| 125 | St. Kitts and Nevis | KNA | 49.2 | 8.4 | -4.2% | 59.6 | 10.5 | -4.1% |
| 126 | St. Lucia | LCA | 36.3 | 12.7 | -2.5% | 46.5 | 14.3 | -2.8% |
| 127 | St. Vincent and the Grenadines | VCT | 59 | 16.6 | -3.0% | 78.0 | 18.3 | -3.5% |
| 128 | Sudan | SDN | 91.3 | 47.6 | -1.6% | 148.3 | 70.1 | -1.8% |
| 129 | Swaziland | SWZ | 101.6 | 44.5 | -2.0% | 146.8 | 60.7 | -2.1% |

| Code | Country Name | Acronym | IMR75 | IMR15 | $g_{IMR}$ | U5MR75 | U5MR15 | $g_{U5MR}$ |
|---|---|---|---|---|---|---|---|---|
| 130 | Sweden | SWE | 8.7 | 2.4 | -3.1% | 10.3 | 3 | -3.0% |
| 131 | Switzerland | CHE | 11.2 | 3.4 | -2.9% | 13.8 | 3.9 | -3.0% |
| 132 | Syrian Arab Rep. | SYR | 59.1 | 11.1 | -4.0% | 79.8 | 12.9 | -4.3% |
| 133 | Tajikistan | TJK | 106.7 | 38.5 | -2.5% | 139.4 | 44.8 | -2.7% |
| 134 | Tanzania | TZA | 118.2 | 35.2 | -2.9% | 198.3 | 48.7 | -3.4% |
| 135 | Thailand | THA | 57.6 | 10.5 | -4.1% | 77.4 | 12.3 | -4.4% |
| 136 | Togo | TGO | 119.2 | 52.3 | -2.0% | 200.2 | 78.4 | -2.3% |
| 137 | Tonga | TON | 30.4 | 14.4 | -1.8% | 37.2 | 16.7 | -1.9% |
| 138 | Trinidad and Tobago | TTO | 39.5 | 18.2 | -1.9% | 46 | 20.4 | -2.0% |
| 139 | Tunisia | TUN | 94.2 | 12.1 | -4.9% | 135.1 | 14 | -5.4% |
| 140 | Turkey | TUR | 108.8 | 11.6 | -5.3% | 158.6 | 13.5 | -5.8% |
| 141 | Tuvalu | TUV | 61.7 | 22.8 | -2.4% | 83.9 | 27.1 | -2.7% |
| 142 | Uganda | UGA | 118.4 | 37.7 | -2.8% | 198.7 | 54.6 | -3.1% |
| 143 | Ukraine | UKR | 25 | 7.7 | -2.8% | 30 | 9 | -2.9% |
| 144 | United Arab Emirates | ARE | 45.9 | 5.9 | -4.9% | 59.4 | 6.8 | -5.1% |
| 145 | United Kingdom | GBR | 15.4 | 3.5 | -3.5% | 18 | 4.2 | -3.5% |
| 146 | United Countrys | USA | 16 | 5.6 | -2.5% | 18.8 | 6.5 | -2.6% |
| 147 | Uruguay | URY | 49.4 | 8.7 | -4.1% | 55.4 | 10.1 | -4.1% |
| 148 | Vanuatu | VUT | 64.6 | 23.1 | -2.5% | 88.5 | 27.5 | -2.8% |
| 149 | Venezuela, RB | VEN | 41.7 | 12.9 | -2.8% | 53.2 | 14.9 | -3.1% |
| 150 | Vietnam | VNM | 51.8 | 17.3 | -2.6% | 77.6 | 21.7 | -3.1% |
| 151 | West Bank and Gaza | PSE | 78 | 18 | -3.5% | 109.5 | 21.1 | -3.9% |
| 152 | Yemen, Rep. | YEM | 181.2 | 33.8 | -4.0% | 271.6 | 41.9 | -4.5% |
| 153 | Zambia | ZMB | 96.3 | 43.3 | -1.9% | 157.2 | 64 | -2.2% |
| 154 | Zimbabwe | ZWE | 70.3 | 46.6 | -1.0% | 110.4 | 70.7 | -1.1% |

The values of the columns 4-5 and 7-8 correspond to the two child mortality rates per 1,000 live births at the beginning of the sample (IMR75 and U5MR75), at the end of the sample (IMR15 and U5MR15), and the average growth rate ($g_{IMR}$ and $g_{U5MR}$) for the total sample (1975-2015).

Table 2.2. Testing for convergence

| | 1975-2015 sample | | 1975-2007 sample | |
|---|---|---|---|---|
| | β | Log-t | β | Log-t |
| IMR | -1.15 | -219.30* | -1.22 | -335.93* |
| U5MR | -1.07 | -168.431* | -1.16 | -507.15* |

The term log-t stands for a parameter which is twice the speed of convergence of this club towards the average. Log-t is the convergence test statistic, which is distributed as a simple one-sided t-test with a critical value of −1.65.
* indicates the rejection of the null hypothesis of convergence.

Table 2.3. Estimated convergence clubs for infant mortality rates

| | Africa | America | Asia | Europe | Oceania |
|---|---|---|---|---|---|
| club1 | CAF TCD LSO MLI NER SLE SWZ ZWE BEN BWA BFA CMR COD CIV GIN LBR MRT MOZ NGA COM | DMA | AFG PAK | | PNG |
| club2 | BDI COG ERI ETH GMB GHA KEN MWI NAM RWA STP SEN ZAF SDN TZA TGO UGA YEM ZMB | BOL GUY HTI | BGD IND IRQ MDG MMR TJK | | KIR MHL SLB VUT |
| club3 | DZA CPV MAR SYC | DOM VCT TTO BRB COL ECU GTM HND JAM NIC PAN PRY LCA | BTN KHM EGY IDN JOR KGZ NPLPHL VNM PSE | | FJI TON TUV |
| club4 | MUS LBY TUN | BLZ VEN ARG BHS BRA CRI SLV MEX PER URY | CHN IRN KAZ KWT MDV RUS SAU LKA SYR THA TUR | BGR GEO MDA ROU UKR | |
| club 5 | LBN | CAN CHL KNA USA | ARE BHR MYS OMN QAT | MLT HUN | NZL |
| club 6 | | CUB | ISR KOR | LTU POL CHE GBR AUT BEL DNK FRA DEU GRC IRL NLD ESP | AUS |
| club 7 | | | | PRT SWE ITA | |
| club 8 | | | JPN. SGP | FIN NOR | |
| club 9 | | | | ISL LUX | |

The convergence clubs reported have been obtained by applying the PS algorithm which aims to find groups of countries with similar convergence speeds to the average. Adjacent clubs have been joined if suggested by the PS statistic.

Table 2.4. Estimated convergence clubs for under-5 mortality rates

| | Africa | America | Asia | Europe | Oceania |
|---|---|---|---|---|---|
| club1 | CAF TCD LSO MLI NER SLE SWZ ZWE | | | | |
| club2 | BEN BWA BFA CMR COD CIV GIN LBR MRT MOZ NGA | | AFG | | |
| club3 | BDI COM COG ERI ETH GMB GHA KEN MWI NAM RWA STP SEN ZAF SDN TZA TGO UGA YEM ZMB | BOL DMA DOM GUY HTI VCT TTO | BGD IND IRQ MDG MMR PAK TJK | | FJI KIR MHL PNG SLB VUT |
| club4 | DZA CPV MUS MAR SYC | BRB BLZ COL ECU GTM HND JAM NIC PAN PRY LCA VEN | BTN KHM EGY IDN JOR KGZ NPL PHL VNM PSE | | TON TUV |
| club 5 | LBY TUN | ARG BHS BRA CRI SLV MEX PER KNA USA URY | CHN IRN KAZ KWT MDV RUS SAU LKA SYR THA TUR | BGR GEO MDA ROU UKR | |
| club 6 | LBN | CAN CHL | BHR MYS OMN QAT ARE | MLT | NZL |
| club 7 | | CUB | | HUN LTU POL CHE GBR | |
| club 8 | | | ISR KOR | AUT BEL DNK FRA DEU GRC IRL NLD PRT ESP | AUS |
| club 9 | | | JPN | FIN NOR SWE | |
| club 10 | | | SGP | ISL | |

The convergence clubs reported have been obtained by applying the PS algorithm which aims to find groups of countries with similar convergence speeds to the average. Adjacent clubs have been joined if suggested by the PS statistic.

Table 2.5. Descriptive analysis of the explanatory variables

| | club 1 | club2 | club3 | club4 | club 5 | club6 | club 7 | club 8 | club 9 |
|---|---|---|---|---|---|---|---|---|---|
| Battle-related deaths | 584.9 | 981.6 | 183.0 | 1,444.7 | 38.5 | 20.8 | 0.0 | 0.0 | 0.0 |
| Presence of peacekeepers | 2,067.5 | 623.0 | 2.4 | 29.9 | 807.2 | 0.0 | 0.0 | 0.0 | 0.0 |
| Intentional homicides | 9.6 | 8.3 | 13.4 | 10.9 | 4.0 | 2.0 | 1.1 | 1.1 | 0.8 |
| Population density | 74.7 | 165.5 | 153.7 | 147.3 | 358.8 | 198.9 | 114.6 | 2,046.8 | 111.6 |
| Urban population | 38.0 | 39.1 | 48.9 | 67.8 | 85.0 | 79.0 | 72.7 | 89.5 | 92.1 |
| Births | 61.5 | 69.7 | 91.0 | 97.6 | 99.0 | 99.0 | 99.6 | 99.6 | 98.8 |
| Total fertility rate | 4.9 | 3.9 | 2.5 | 2.0 | 1.9 | 1.7 | 1.5 | 1.5 | 1.6 |
| Basic sanitation services | 29.0 | 42.3 | 80.4 | 90.8 | 99.1 | 97.9 | 99.3 | 99.4 | 98.2 |
| Per capita GDP | 1,191 | 1,799 | 4,886 | 9,527 | 30,190 | 39,728 | 37,007 | 58,453 | 76,444 |
| Total unemployment rate | 7.2 | 7.4 | 7.3 | 7.5 | 4.4 | 8.4 | 10.6 | 4.7 | 5.35 |
| Private health expenditure | 3.4 | 3.1 | 2.9 | 2.7 | 2.7 | 2.4 | 2.5 | 2.1 | 1.4 |
| Public health expenditure | 9.1 | 10.4 | 11.8 | 12.2 | 12.8 | 15.6 | 14.9 | 16.3 | 14.65 |
| Total health expenditure | 5.9 | 6.0 | 6.4 | 6.5 | 7.4 | 9.4 | 10.2 | 8.6 | 7.9 |
| Female literacy rate | 42.9 | 63.5 | 80.7 | 87.3 | 90.1 | 97.2 | 94.1 | 96.1 | 98.0 |

Average values of the different explanatory variables when the countries are classified according to the estimated convergence clubs for IMR (see Table 2.2).

Table 2.6. Ordered probit estimates of IMR and U5MR convergence clubs

| | IMR | Elasticity | U5MR | Elasticity |
|---|---|---|---|---|
| Population density | $2.9 \times 10^{-4}$ (2.81) | -0.4 | $4.9 \times 10^{-4}$ (3.40) | -0.9 |
| Total fertility rate | -0.816 (-4.48) | 7.8 | -0.675 (-4.01) | 8.2 |
| Per capita GDP | $1.8 \times 10^{-5}$ (2.30) | -2.2 | $8.8 \times 10^{-6}$ (1.05) | -1.1 |
| Basic sanitary services | 0.043 (5.51) | -18.2 | 0.035 (4.67) | -16.5 |
| Urban population | 0.022 (3.68) | -7.2 | 0.026 (4.09) | -9.9 |
| Total unemployment rate | -0.076 (-3.66) | 2.5 | -0.060 (-2.54) | 2.3 |
| Public health expenditure | 0.049 (2.26) | -3.1 | 0.40 (1.78) | -2.9 |
| $D_{AFRICA}$ | -0.986 (-2.26) | | -1.449 (-3.37) | |
| $D_{AMERICA}$ | -1.674 (-4.47) | | -1.896 (-5.07) | |
| $D_{OCEANIA}$ | -1.675 (-4.07) | | -1.468 (-2.64) | |
| $D_{ASIA}$ | -1.064 (-3.10) | | -0.844 (-2.63) | |
| $D_{OECD}$ | 1.676 (3.39) | | 1.881 (4.58) | |
| Cutpoints | | | | |
| Cut1 | -2.527 | | -3.672 | |
| Cut2 | 0.083 | | -2.584 | |
| Cut3 | 2.127 | | 0.381 | |
| Cut4 | 4.405 | | 2.089 | |
| Cut5 | 5.744 | | 4.224 | |
| Cut6 | 7.733 | | 5.302 | |
| Cut7 | 8.155 | | 6.058 | |
| Cut8 | 9.099 | | 7.658 | |
| Cut9 | | | 8.644 | |
| Pseudo R2 | 0.57 | | 0.53 | |

The dependent variable takes the value i when country is included in club i, with i=1, 2, …, M and M being 9 for IMR and 10 for U5MR.

The values in parenthesis are the robust t-statistics for testing the null hypothesis whose associated coefficient is 0.

Columns 2 and 4 present the estimated coefficients, whilst columns 3 and 5 reflect the corresponding estimated elasticities.

Figure 2.1. Evolution of the child mortality rates. 1975-2015.

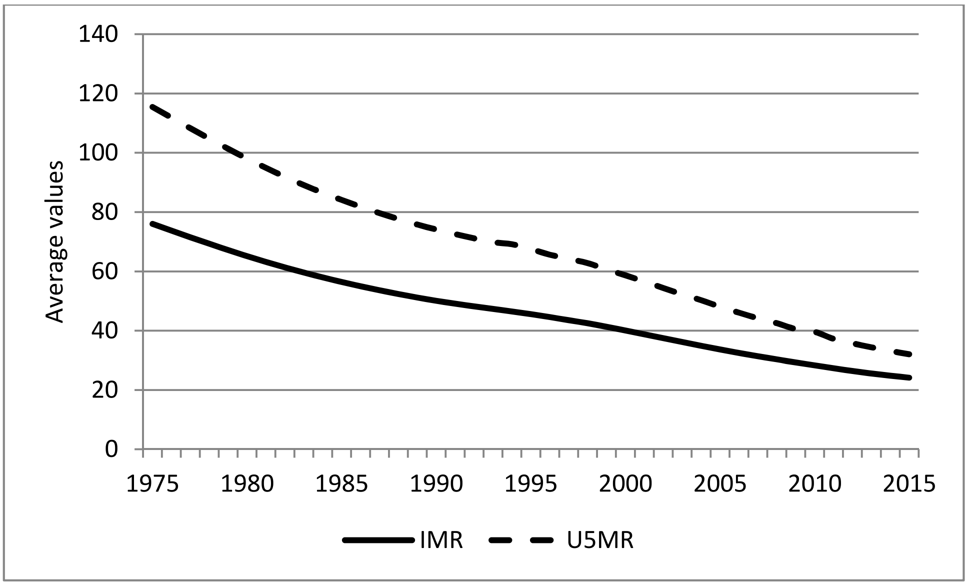
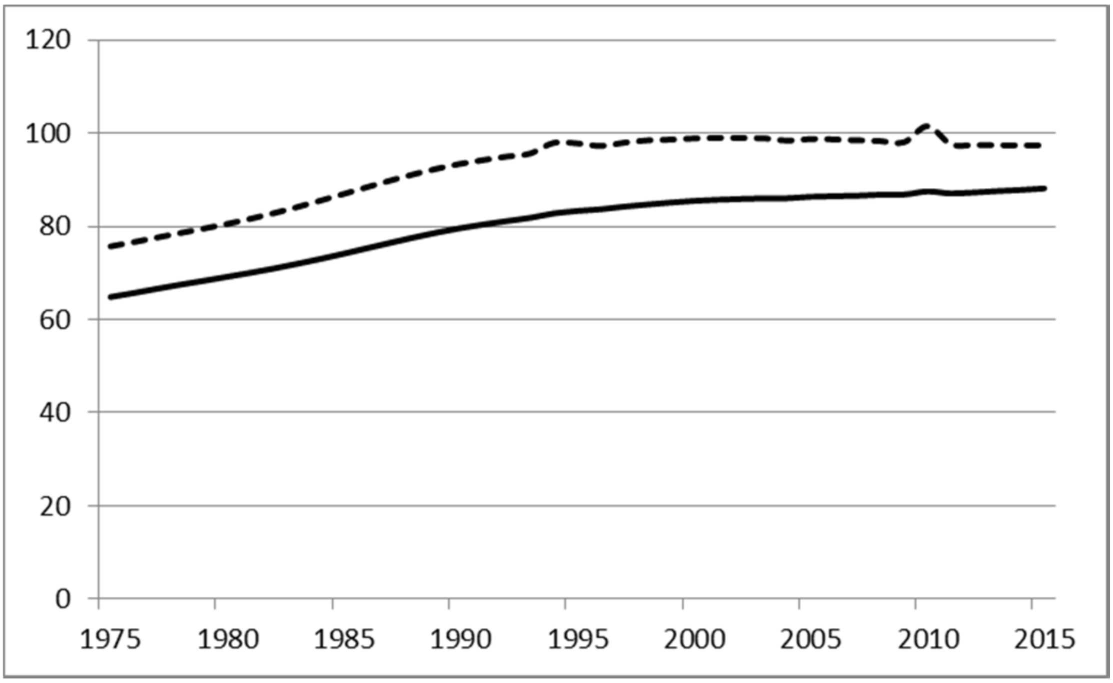
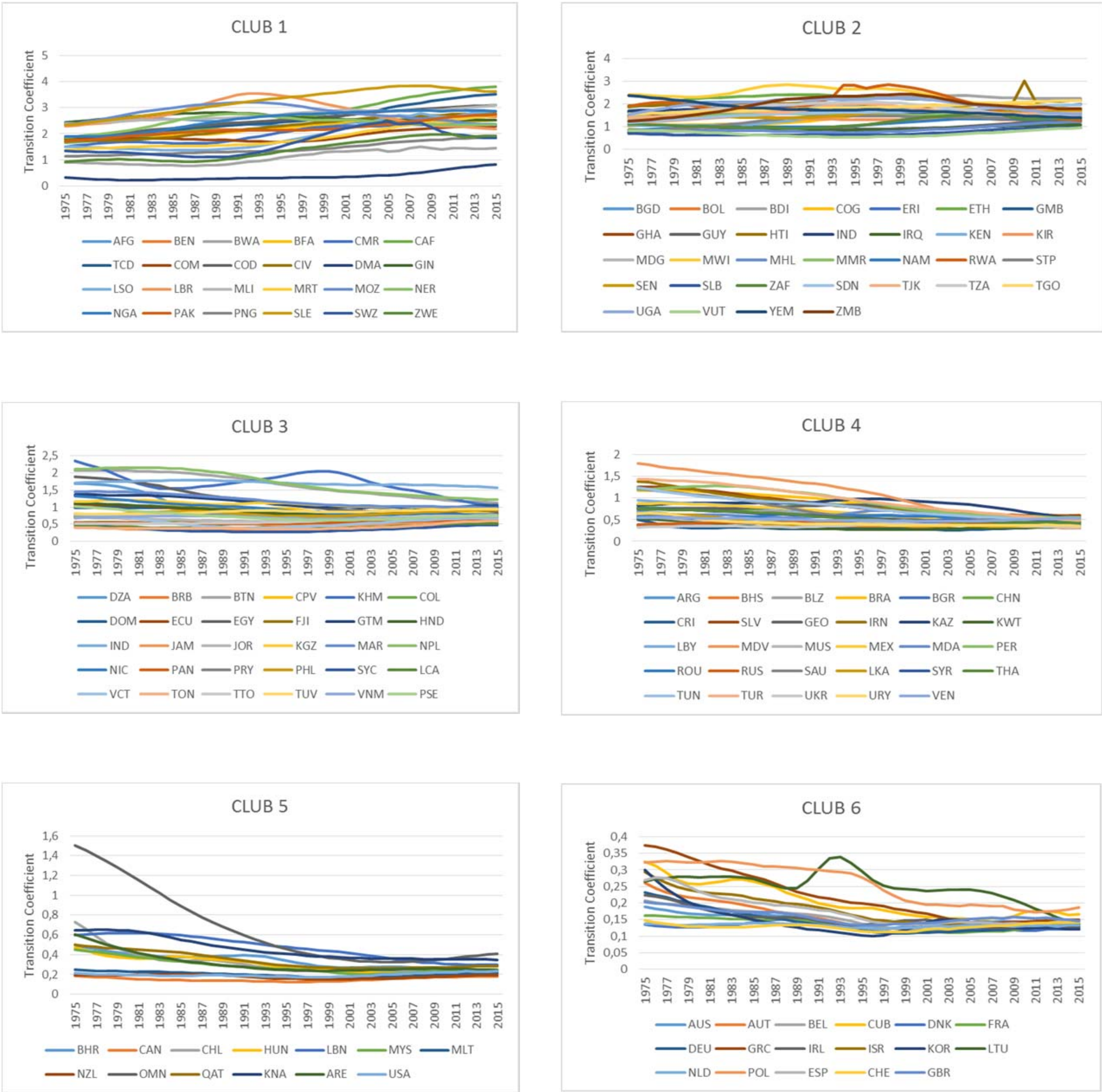


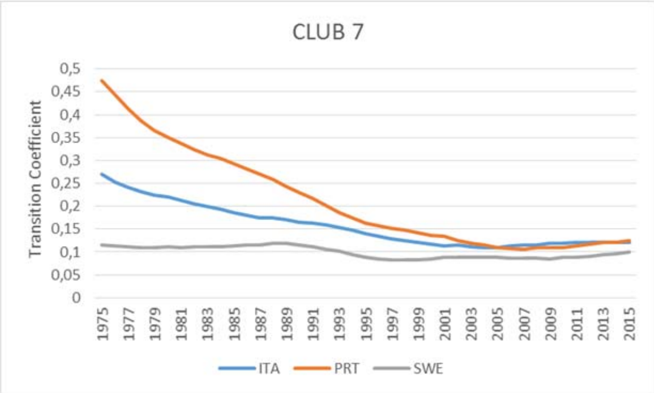
Figure 2.2. σ -convergence analysis for child mortality

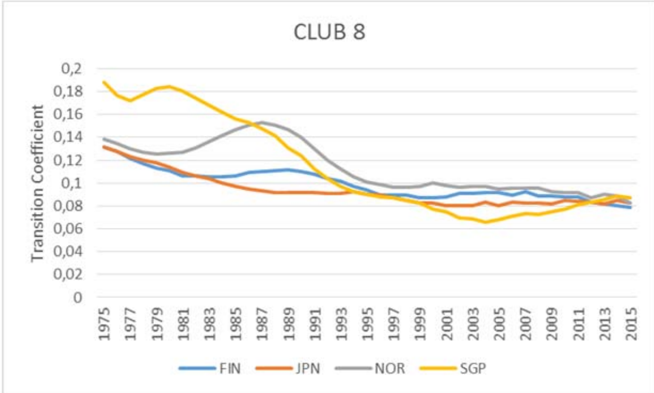


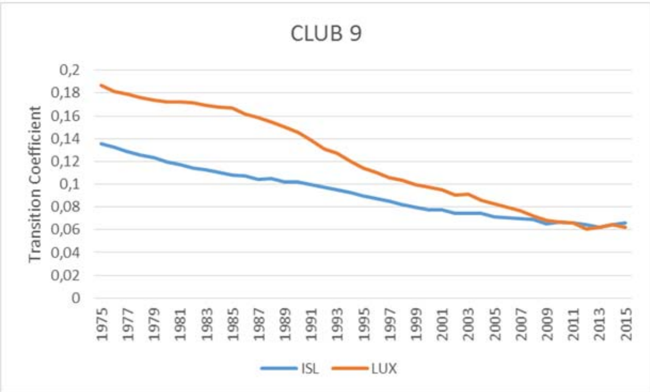

This figure reflects the σ-convergence analysis by presenting the values of the variation coefficient of the U5MR (dotted line) and IMR (solid line).

Figure 2.3. Transition coefficient evolution ($h_{it}$) of IMR.



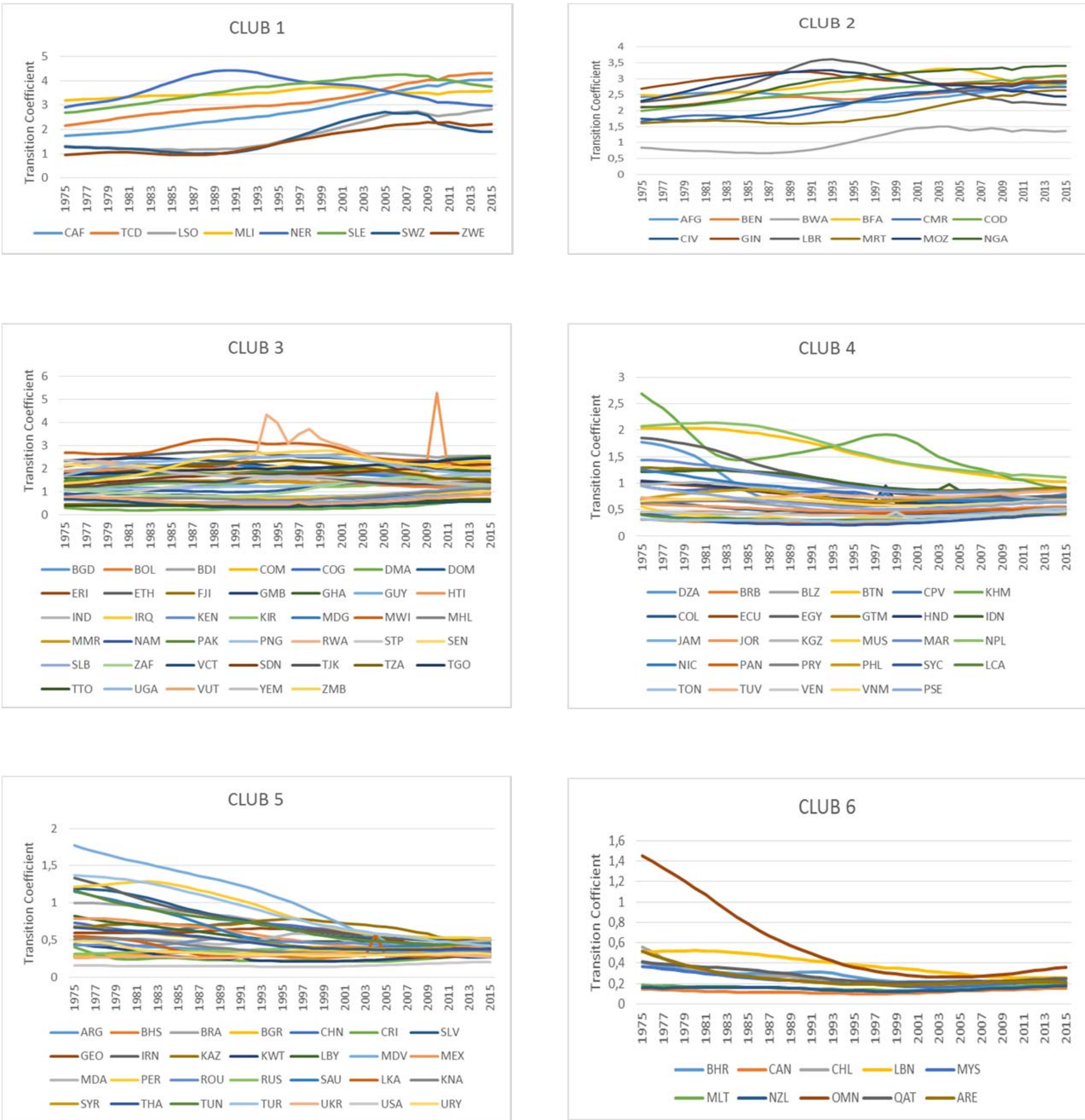

CLUB 7
Transition Coefficient
ITA
PRT
SWE

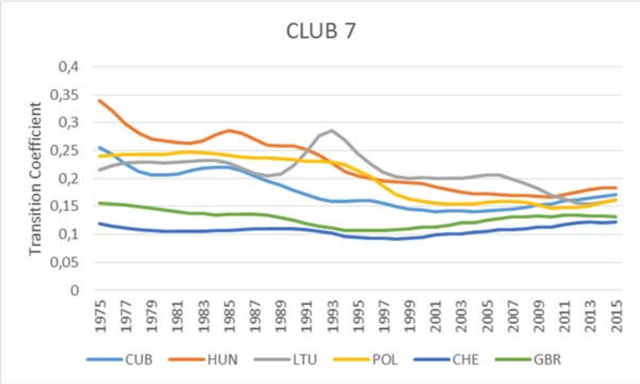

CLUB 8
Transition Coefficient
FIN
JPN
NOR
SGP

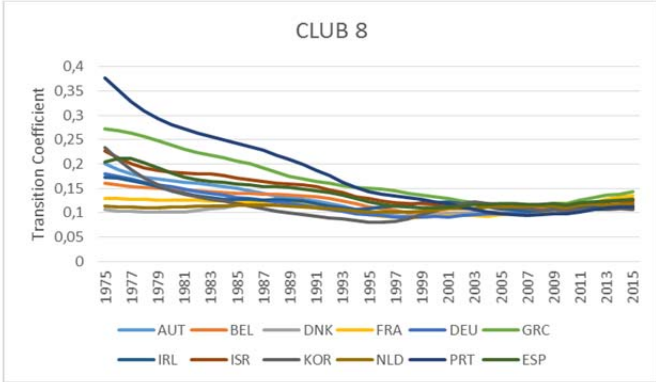

CLUB 9
Transition Coefficient
ISL
LUX

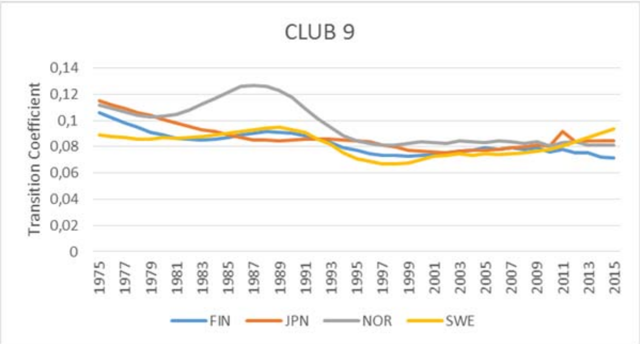
.

Figure 2.4. Transition coefficient evolution ($h_{it}$) of U5MR.




CLUB 7
Transition Coefficient
CUB
HUN
LTU
POL
CHE
GBR


CLUB 8
Transition Coefficient
AUT
BEL
DNK
FRA
DEU
GRC
IRL
ISR
KOR
NLD
PRT
ESP


CLUB 9
Transition Coefficient
FIN
JPN
NOR
SWE

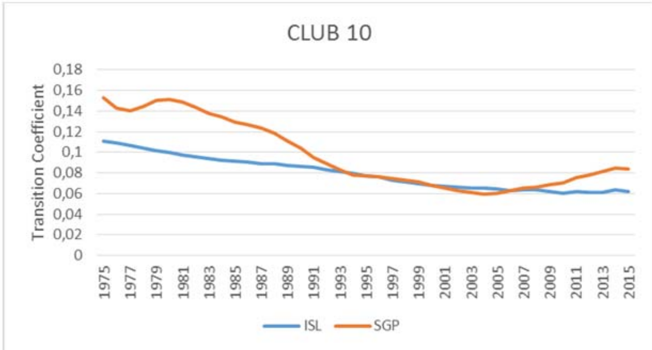

CLUB 10
Transition Coefficient
ISL
SGP

Figure 2.5. Child mortality rates means for convergence clubs

Figure 2.5.a. Arithmetic mean of IMR for each convergence club.

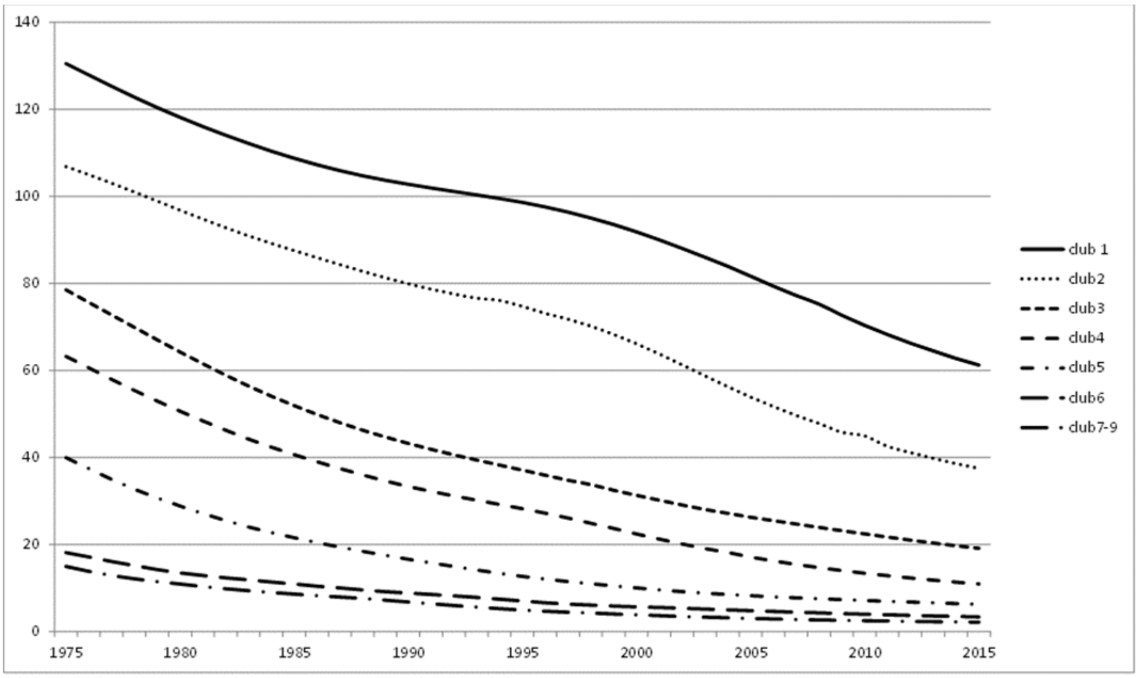


Figure 2.5.b. Arithmetic mean of U5MR for each convergence club.

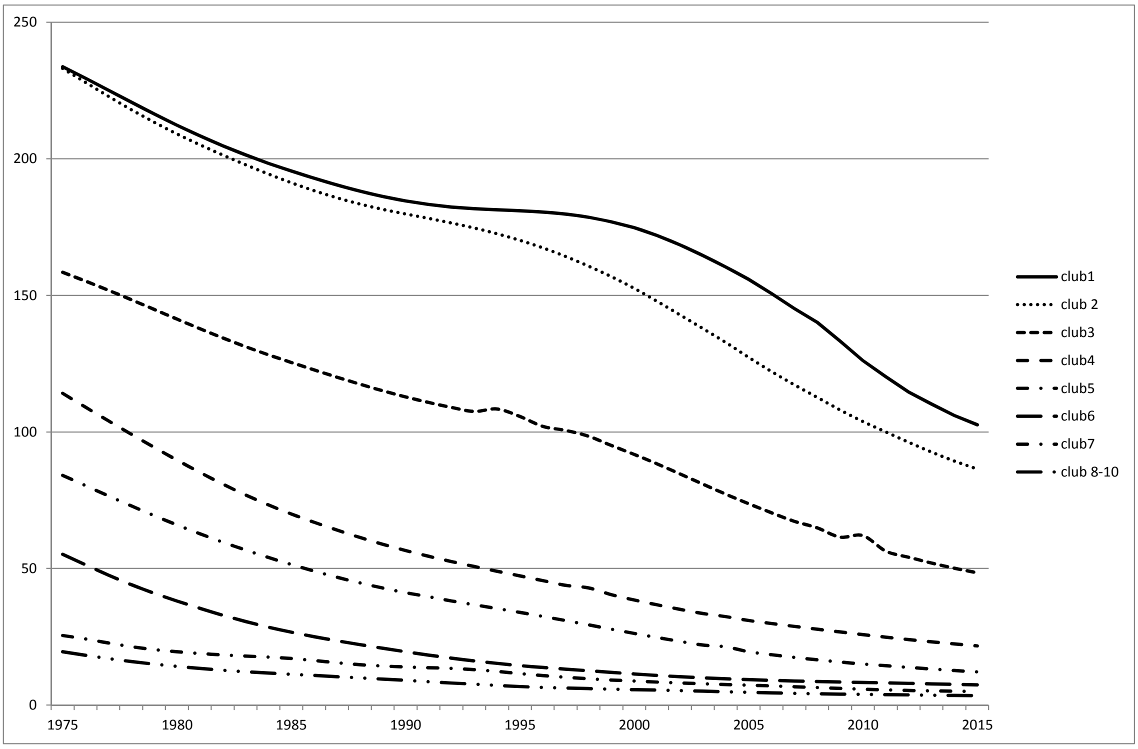
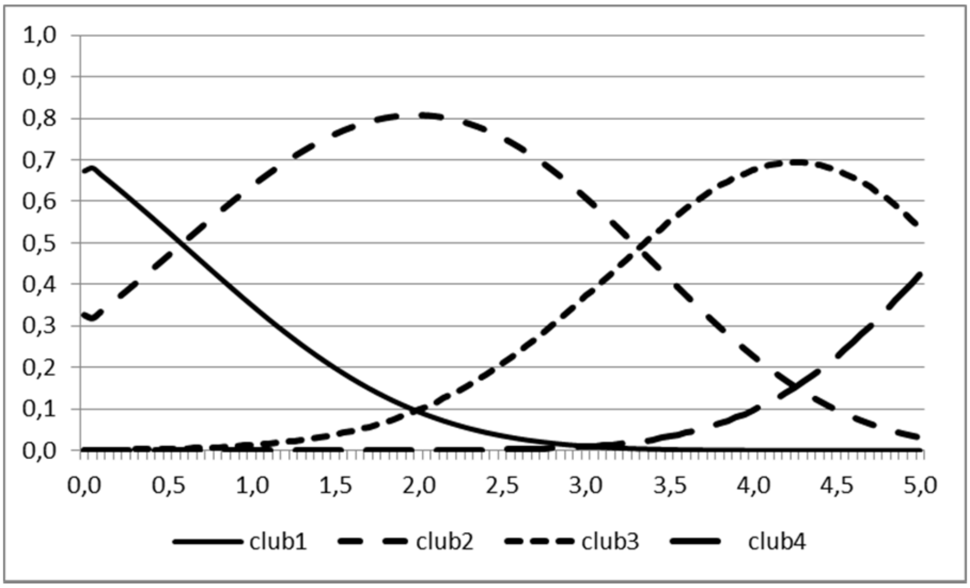


Figure 2.6. Evolution of the estimated probabilities for different growth rates.



# Chapter 3.

# Convergence in infant mortality rates: The Spanish regions case

## 3.1. Introduction

The infant mortality rate (IMR) is one of the main health indicators collected by various international organizations. In order to understand the importance of this indicator, we should simply note that the United Nations explicitly included the reduction of the child mortality rate as a target of the Millennium Development Goals (MDGs), seeking a reduction of two thirds of the child mortality rate in 2015. Likewise, this organization has again emphasized the role of the fight against child and maternal mortality as a very important part of its Sustainable Development Goals (SDGs), with the aim of eradicating preventable deaths of newborns and children under 5 by 2030.

The efforts made to achieve these goals have been very fruitful and, although some Sub-Saharan and south Asian zones still exhibit large values, the reduction in IMR has been remarkable. The global under-5 mortality rate has declined by more than half in 1990-2015, whilst child deaths have fallen from 90 to 43 deaths per 1000 live births in this same period. This reduction has not affected to all the countries in the same way. In a recent paper, Martínez-Asenjo and Montañés (2017) prove that the worldwide child mortality rates do not converge, rather we observe the presence of several patterns of behavior. This would not be very surprising, given that the sample includes very different countries. These authors find statistical differences even between the European countries, which can be grouped into different convergence clubs. This result is additionally supported by some previous analysis. For instance, Gil-Alaña et al. (2017) find important differences in the time series properties of the child mortality rates of the OECD countries, which is incompatible with the convergence of these rates.

However, the interpretation of the results obtained from an international comparison of child mortality rates is often handicapped by the noticeable differences between the health policies and health systems of these countries. For instance, it is not easy to compare the results of Spain, Italy and Portugal, whose health policies are based on National Health Care systems, with the ones of France and Germany, whose Health Care system is based on a social insurance system. But, we can also find disparities inside these two groups. Therefore, it is possible that the differences in the reduction of the IMR have been caused by the disparity of health systems.

We would expect that the regional analysis of the IMR of a given country should offer more evidence of convergence. In this regard, Sidebotham et al. (2014), who study the case of the England and Wales, and Latto (2015), who analyses the Finnish case, cast some doubts on this expectation and offer evidence of regional differences in the IMR. Something similar occurs for the case of Spain, the one we want to consider here. Given that all the regions have a quite similar health system, the IMR of the Spanish regions should be similar. However, this does not seem to be the case. The seminal works of Regidor et al. (1994, 1995) and Regidor et al. (2002) find remarkable differences in both infant and adult mortality patterns of behavior across the Spanish regions.

Moreover, these authors also offer evidence of the importance of socio-economics factors to explain these disparities.

In spite of the unquestionable interest of the above mentioned papers, we consider that they do not totally exploit the time series properties of the variables. Consequently, their results are not time consistent. For instance, the presence of trends in the IMR could lead us to observe some very interesting phenomena, such as the possible existence of stochastic convergence. This concept is commonly employed in economics to explain the growth of economies. We should note that Barro and Sala-i-Martin (1991, 1992) introduced the notion of convergence to assess whether the poor regions (or countries) grow faster than the rich ones, implying that they will catch up (β-convergence) in the long-run, or whether the dispersion of income diminishes over time (σ-convergence). We can adapt both concepts to the case of IMR and determining whether the regions with the highest infant mortality rates have converged towards the ones with the lowest rates. Furthermore, the recent advances in this type of studies allow us to determine the presence of multiple patterns of behavior. So, the regions can be grouped into different convergence clubs and, subsequently, we can try to explain what determines membership to them.

Following this background, we intend to answer the next questions:

- Does convergence in the IMR exist between Spanish regions?
- Has the great economic recession influenced the regional differences of IMR in Spain?
- What have been the determining factors in the evolution of the IMR in Spain?

To do this we use time series tools to analyze the behavior of the Spanish regional IMR, paying special attention to the study of the convergence hypothesis, which could provide us information about the existence of regional differences in this rate. To that end, the rest of the paper is organized as follows. Section 3.2 presents the data base that will be employed. The methodology for testing for the null hypothesis of convergence is reported in section 3.3, along with the results of its application. Section 3.4 tries to find the factors that can help us to explain the disparities previously found. The paper ends with the main conclusions.

### 3.2. Database description

The regional IMR have been provided by the Spanish Institute of Statistics. This variable is defined as the ratio of the number of deaths in the first year of life over the number of live births occurring in the same population during the same period of time. The data are available for the 17 Spanish regions and cover 1975-2017 period. Table 3.1 includes the acronyms of these regions and several descriptive statistics.

If we consider the evolution of the total Spanish IMR, we can observe that it has declined from its maximum 18.9 in 1975 to its minimum 2.7 in 2017, following the pattern of behavior of the developed countries. However, differences can be seen if we split the sample. The growth rate is -5.6% for 1975-1999, whilst it is -2.7% for 2000-2017. The rate slightly declines (in absolute terms) to -2.6% for 2000-2012, when the Spanish government took drastic decisions which implied serious cuts in public expenditures, including health services. After these economic decisions, we can observe that the average growth rate of 2013-2017 is no longer negative. Thus, the cuts in health seem to have negatively affected the evolution of the Spanish infant mortality rates, at least from an aggregate point of view.

If we now turn to the evolution of the Spanish regions, we observe that CYL presents the highest IMR in 1975 (24.0), 5 points above the Spanish rate (18.9). By contrast, MAD reflects the lowest rate at the beginning of the sample (14.8), just 4 points below the national rate. At the end of the sample, CYL again presents the highest rate (3.4), whilst NAV now exhibits the lowest rate (2.1). We can also observe that the distance between the regional IMR reduces.

The clear reduction in the rates could be compatible with the presence of convergence, in the sense that, if all the regions presented similar IMR at the end of the sample, then the variance would be reduced, suggesting evidence of σ-convergence. To verify this, the evolution of the coefficient of variation across the sample is graphically represented in Figure 3.1. It is true that we can observe an initial reduction in the coefficient of variation for 1975-1983 but it later begins growing until 2012, when the trend changes and the coefficient of variation diminishes up to the values of the 1980's. Therefore, the analysis of this figure would not lead us to think that the Spanish IMR have converged. Rather, this figure leads us to consider that they may have diverged. However, this previous analysis is far from conclusive and, consequently, we should employ more powerful tools in order to test for convergence, which is the aim of the next section.

### 3.3. Testing for convergence in infant mortality

The previous analysis has shown the disparities that exist between the IMR of the Spanish regions, which cast serious doubts on the existence of convergence. We now formally test for the convergence hypothesis. Following the recent papers of Phillips and Sul (2007, 2009) (PS hereafter) where they develop a framework that allows us, first, to test for the convergence hypothesis and, if this hypothesis is rejected, to estimate the number of convergence clubs. The PS methodology is closely related to the standard σ-convergence analysis, a concept introduced by Barro and Sala-i-Martin (1990) and Barro et al. (1991), in that it tests for the decline of the variable of interest over time in the cross-sectional dispersion. However, it clearly outperforms the classical convergence analysis, namely, the abovementioned σ-convergence concept and the β-convergence concept, which was first introduced by Baumol (1986). The PS methodology is based on a general nonlinear time-varying factor model, which admits the presence of transitional heterogeneity. Additionally, it is flexible with respect to the time properties of the variables under analysis because it does not impose any particular assumption about them. Finally, it is clearly free of the criticism received by the β-convergence analysis in De Long (1988) and Quah (1993). As a consequence, the use of this methodology has recently become very popular in convergence analysis.

Following these authors, let us consider that $X_{it}$ represents the IMR considered in this paper, with i=1, 2, …, 17 (the 17 Spanish regions) and t= 1975, …, 2017. This variable can be decomposed as $X_{it} = \delta_{it}\,\mu_t$, where $\mu_t$ and $\delta_{it}$ are the common and the idiosyncratic component, respectively. PS suggest testing for convergence by analyzing whether $\delta_{it}$ converges towards δ. To do so, they first define the relative transition component or transition coefficient:

$$h_{it} = \frac{X_{it}}{N^{-1}\sum_{i=1}^{N} X_{it}} = \frac{\delta_{it}}{N^{-1}\sum_{i=1}^{N} \delta_{it}} \qquad (3.1)$$

In the presence of convergence, $h_{it}$ should converge towards unity, while its cross-sectional variance, $H_{it}$, is defined as follows:

$$H_{it} = N^{-1}\sum_{i=1}^{N}(h_{it} - 1)^2 \qquad (3.2)$$

and should go to 0 when T goes towards infinity. Then, PS test for convergence by estimating the following equation:

$$log\frac{H_1}{H_t} - 2log[\log(t)] = \alpha + \beta\log(t) + u_t,\ t = T_o, \dots, T \quad (3.3)$$

with $T_o = [rT]$, and r=0.3, as suggested by PS. Equation (3.3) is commonly known as the log-t regression. The null hypothesis of convergence is rejected whenever parameter β is lower than 0. PS suggest estimating model (3.3) by methods which correct for the presence of autocorrelation and heteroscedasticity and, later, employ the t-statistic to test the null hypothesis β=0. The use of these robust methods ensures that this t-ratio converges towards a standard N(0,1) distribution and, therefore, we will reject the null hypothesis of convergence whenever this t-statistic takes values lower than -1.65.

If we reject convergence, PS propose the following robust clustering algorithm for identifying clubs in a panel of countries:

i. Order the N regions according to their final values.

ii. Starting from the highest-order region, add adjacent regions from our ordered list and estimate model (3.3). Then, select the core group by maximizing the value of the convergence t-statistic, subject to the restriction that it is greater than -1.65.

iii. Continue adding one region at a time of the remaining regions to the core group, and re-estimate model (3.3) for each formation. Use the sign criterion (t-statistic >0) to decide whether a region should join the core group.

iv. For the remaining regions, repeat steps (ii)–(iii) iteratively and stop when clubs can no longer be formed. If the last group does not have a convergence pattern, conclude that its members diverge.

PS recommend performing club merging tests after running the algorithm using equation (3.3) in order to avoid the over-estimation of the number of clubs.

Finally, we have followed the suggestion of PS and we have extracted the trend components of the series by filtering them using the Hodrick and Prescott (1997) filter, applying the standard value $\lambda$=400.

The result that we have obtained are presented in Panel A of Table 3.2. This table reflects the results for the total IMR. We have also considered the pre Great Recession sample (1975-2007). We can first observe that the null hypothesis of convergence is rejected for both samples. This leads us to conclude that the IMR of the Spanish regions do not follow a single pattern of behavior. Rather, they diverge and, probably, different patterns of behavior may co-exist. Furthermore, if we compare the estimated values of the parameter $\beta$ for the two samples, we can see that they are clearly different, the estimation of $\beta$ for the pre Great Recession being approximately multiply by 4 the estimated value for the total sample. This would imply that the evidence against the null hypothesis was greater for 1975-2007 than for the total sample, inducing us to think that the cuts in health services since 2008 have affected the evolution of the IMR. In order to study this question more deeply, we have estimated the $\beta$ parameter for the samples that cover the period 1975-N, with N = 2000, 2001, … 2017. Figure 3.2 reflects the evolution of the estimation of the parameter $\beta$. We can observe that it was between -2.4 and -3.1 until 2007, whilst this value increased during 2008-2017, especially since 2012. This result implies that the evidence against the null hypothesis of convergence has reduced during the Great Recession period. Therefore, the policies that aimed to reduce public expenditure, especially in the health services, have affected the convergence process, diminishing the disparities in IMR between the Spanish regions. Moreover, if we assume that the recent trend will continue, we could expect these disparities to disappear in the near future. However, the convergence direction is not clear and it is possible that the regions with the lowest IMR become closer to those with the largest ones. If this is the case, it could be interpreted as an adverse effect of the recent health cuts. We will back to this point later.

Once we have rejected the convergence null hypothesis, we should consider the possible presence of different convergence clubs. To that end, we can employ the PS algorithm. The results we have obtained, after testing whether the adjacent clubs can be joined, are presented in Panel B of Table 3.2.

The analysis of this table permits us to observe the presence of different patterns of behavior in the IMR of the Spanish regions, finding two different clubs. Figure 3.3 maps these clubs in different colors. We can see the evolution of the different clubs by generating an index that takes the average values of the IMR of the regions included in the different clubs. This is presented in Figure 3.4. The values of the club 2 are the greatest until 1999. Since then, the club 1 commonly exhibits the greatest values, especially since 2010. But, the differences are important for 2013-2017. The club 1 shows an average growth rate of 1.7%, while the values of club 2 slightly vary. This result again warns us about the possible negative effect of the cuts on health expenditure on the evolution of the IMR in Spain.

### 3.4. Which are the forces that drive the club creation?

The previous section has shown that the IMR of the Spanish regions do not follow a single pattern of behaviour across the sample considered. Rather, we can observe multiple patterns. The next step is to investigate the sources of these differences. We should bear in mind that many factors have been used to try to explain them. For instance, some economic factors are commonly employed in order to explain the differences in child mortality. The Gross Domestic Product (GDP) has frequently been employed to explain the evolution of child mortality, as we can see in the papers of Gbesemete and Jonsson (1993), Alves and Belluzzo (2004), Hakobyan et al. (2006), Renton et al. (2012) and Subramaniam et al. (2016). Ko et al. (2014) also consider the labour market as a possible explanatory variable.

From a different perspective, some authors use educational levels as another important factor to understand the evolution of mortality. We can cite the papers of Mondal et al. (2009), Defo (1996), Matsaganis (1992), and Bourne and Walker (1991). Some authors have also found an interesting relationship between child mortality and female fertility. In particular, some recent analyses have found a negative relationship between these two variables, as is the case of Hondroyiannis and Papapetrou (2002) for Finland and Narayan and Peng (2007) for Japan.

Demographic factors should also be taken into account. There is a large body of literature which associates health with place of residence. For instance, Jankowska et al. (2014), Fink et al. (2014) and Quansah et al. (2016) have recently found evidence of the comparative advantages of urban over rural residence. This is mainly due to the improvement in living standards offered by urban zones, namely, access to better health care services, education and housing. However, we should note that the influence of urban population on IMR seems to weaken in recent years, if we consider the results of Kimani-Murage et al. (2014). Hathi et al. (2014) also find a relationship between population density and infant mortality.

Finally, we should consider that some authors, as Hanmer et al. (2003), have found a strong relationship between child mortality and access to healthcare.

In order to reflect all these potential explanatory factors, the variables that we have selected are presented in Table 3.3 and their data sources in Appendix C. The socioeconomic variables selected are per capita GDP, the total unemployment rate; the health factors are the percentage of the female population that admits smoking daily, the percentage of female population with a body mass index greater than 30 and the percentage of the female population that recognizes to have drunk alcoholic beverages in the last two weeks; the demographic factors are the fertility rate and the population density. We have also considered the terrain roughness, the per capita public health expenditure, the private medical assurance rate and the migration rate.

We have estimated an ordered probit model. The dependent variable is constructed in such a manner that it takes the value i if the region has been assigned to the i-th club, excluding the non-convergent regions. The final specification has been selected by using a general-to-particular strategy where the non-significant variables have been iteratively removed. The results are presented in Table 3.4.

The estimated model includes different factors. The most important in order to explain the creation of the estimated clubs is the alcohol consumption. The higher the percentage of female people in a region that have drunk alcoholic beverages in the last two weeks, the higher the probability of this region to be included in club 1 and, consequently, the higher the IMR. The second factor is the population density, although its marginal effect is lower than for the rest of the variables. Finally, the economic variables also play an important role in the model. The total unemployment rate is included in the model, with this variable commonly used to reflect the cyclical component of an economy. So, according to our results, the higher the total unemployment rate of a region, the more probable the region is included in club 1. This fact is important, given that it connects the economic situation to the evolution of the infant mortality rates. Then, it seems to be sensible to admit that those regions more influenced by the Great Recessions have had to adopt very serious cuts in public services, in general, and in health expenditure, in particular. Assuming this, then the results presented in Table 3.4 would again link the cuts in health expenditure with the increase of the IMR across the Spanish regions.

### 3.5. Conclusions

This paper analyses the evolution of the Spanish regional IMR by testing for the null hypothesis of convergence. Our results lead us to reject this hypothesis, which implies that these mortality rates do not share a similar pattern of behavior. Rather, we can observe the presence of two different convergence clubs.

We have also considered the possible effect of the cuts in health services derived from the Great Recession on the evolution of the IMR. Whilst it is true that the convergence null hypothesis is rejected even when the sample only considers the pre Great Recession period, we can observe that the increment of the sample reduces the evidence against the null hypothesis. Thus, there seems to be a connection between economic decisions and the evolution of the IMR. Moreover, if the present trend continues, we could expect convergence in the near future. But, this convergence process would go in the opposite direction, given that the regions with the lowest IMR would increase their rates, catching-up the regions with the highest IMR.

Finally, we have estimated a model in order to know which forces drive the creation of the convergence clubs. Using an ordered probit, we observe that the probability of a region being assigned to club 1, the one with the highest IMR, grows with the number of female people who admits having drunk alcoholic beverages in the last two weeks, the population density and the total unemployment rate. The inclusion of this last variable could again suggest that cuts in health services and the evolution of the IMR are related. However, this last result should be interpreted with some caution and additional analysis should be carried out, especially an analysis of causality between these two variables.

Table 3.1. Descriptive analysis of the Spanish regional IMR

| Region | Acronym | $IMR_{75}$ | $IMR_{17}$ | $g_{7517}$ | $g_{7599}$ | $g_{0017}$ | $g_{1317}$ |
|---|---|---|---|---|---|---|---|
| Andalucía | AND | 21.7 | 3.1 | -4.4% | -5.6% | -2.5% | 0.0% |
| Aragón | ARA | 15.8 | 2.4 | -4.3% | -3.9% | -5.2% | 5.8% |
| Asturias | AST | 20.3 | 3.2 | -4.2% | -6.5% | -0.5% | 13.9% |
| I. Baleares | BAL | 16.5 | 3.1 | -3.8% | -5.2% | -2.4% | 2.5% |
| Canarias | CAN | 20.1 | 3.1 | -4.2% | -5.5% | -4.2% | 4.4% |
| Cantabria | CAB | 20.4 | 2.7 | -4.6% | -8.9% | 1.5% | -4.2% |
| Castilla y León | CYL | 24.0 | 3.4 | -4.4% | -6.9% | -1.0% | 1.7% |
| Castilla - La Mancha | CLM | 21.0 | 2.7 | -4.6% | -6.8% | -2.2% | 3.0% |
| Cataluña | CAT | 15.3 | 2.4 | -4.2% | -4.8% | -2.1% | -0.6% |
| C. Valenciana | VAL | 16.3 | 2.7 | -4.1% | -5.5% | -1.6% | 2.4% |
| Extremadura | EXT | 22.9 | 3.2 | -4.5% | -6.2% | -2.5% | 9.5% |
| Galicia | GAL | 23.6 | 2.2 | -5.4% | -8.6% | -3.8% | 0.9% |
| Madrid | MAD | 14.8 | 2.4 | -4.1% | -5.4% | -3.1% | -3.6% |
| Murcia | MUR | 23.4 | 2.8 | -4.8% | -4.5% | -5.1% | -4.6% |
| Navarra | NAV | 18.8 | 2.1 | -5.0% | -8.7% | -3.7% | -3.3% |
| País Vasco | PAC | 18.6 | 2.5 | -4.6% | -5.2% | -1.3% | 1.3% |
| La Rioja | LAR | 17.0 | 2.7 | -4.2% | -4.0% | -2.0% | -2.5% |
| **Total Spain** | **SPA** | **18.9** | **2.7** | **-4.4%** | **-5.6%** | **-2.7%** | **0.3%** |

Columns 1 and 2 reflected the name of the Spanish regions and their respective acronyms.

Columns 3 and 4 present the values of IMR for the beginning and the end of the sample. IMR is defined as the ratio of the number of deaths in the first year of life over the number of live births occurring in the same population during the same period of time, across the total sample (1975-2017).

Columns 5-8 reflect the values of $g_{xxyy}$, with $g_{xxyy}$ being the average growth rate of the XX-YY period.

Table 3.2. Testing for convergence.

| Panel A. Testing for convergence | |
|---|---|
| Total sample (1975-2017) | -3.80[a] |
| Pre Great Recession sample (1995-2007) | -22.26[a] |
| **Panel B. Estimated convergence clubs** | |
| *Club 1* | |
| AND, ARA, AST, BAL, CAN, CAB, CYL, CAT, CVA, EXT, MAD, MUR, PAV | |
| *Club 2* | |
| CLM, GAL, LAR | |
| *Divergent Region* | |
| NAV | |

Panel A shows the results of the PS log-t ratio employed for testing the null hypothesis of convergence. Panel B displays the members of each estimated club.

[a] indicates the rejection of the null hypothesis of convergence at a 5% significance level.

Table 3.3. Average values of the explanatory variables for the ordered probit estimation.

| | Club 1 | Club 2 |
|---|---|---|
| Terrain Ruggedness | 193.6 | 154.1 |
| Population density | 168.1 | 56.8 |
| GDP (pc) | 19,393 | 17,765 |
| Unemployment rate | 17.9 | 16.5 |
| Migration rate | 7.5 | 5.7 |
| Public Health Exp. (pc) | 560.7 | 544.1 |
| Private medical assurance rate | 0.7 | 0.4 |
| Female Obesity rate | 17.6 | 19.8 |
| Female Tobacco rate | 20.3 | 19.0 |
| Female alcoholic beverage rate | 38.8 | 35.5 |
| Fertility global rate | 46.9 | 47.7 |

Table 3.4. Estimation of the ordered probit model

| | IMR |
|---|---|
| Female alcoholic bev. consumption | -0.23<br>(-1.99) |
| Population density | -0.006<br>(-1.94) |
| Total Unemployment rate | -0.46<br>(-1.86) |
| Pseudo R2 | 0.55 |
| Number of cases correctly predicted | 14 (82.4%) |

The dependent variable is constructed by assigning the value i to the regions in club i, with i=1,2. The values in parenthesis are the robust t-ratios for testing the single significance of the parameters.

Figure 3.1. σ-convergence analysis for infant mortality rates in Spanish regions.

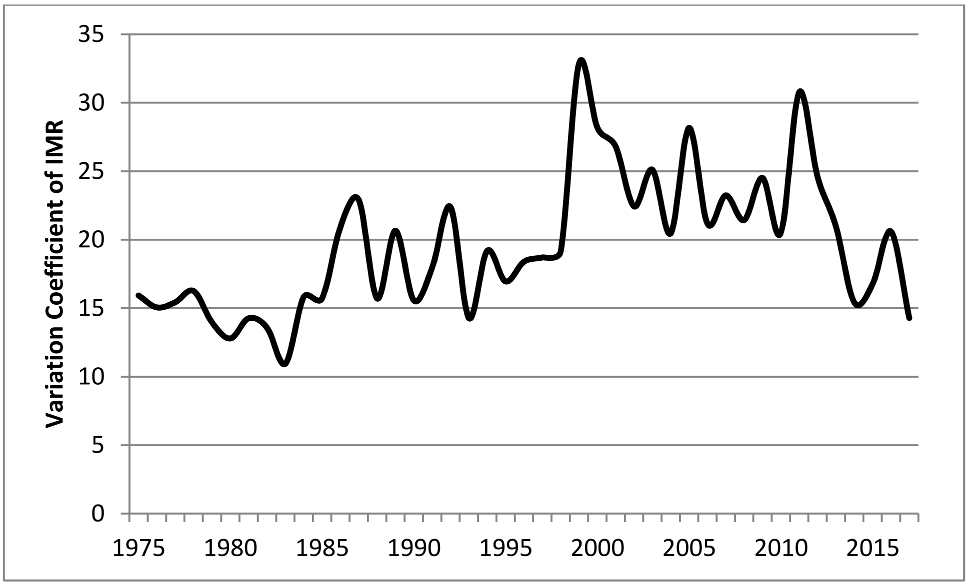
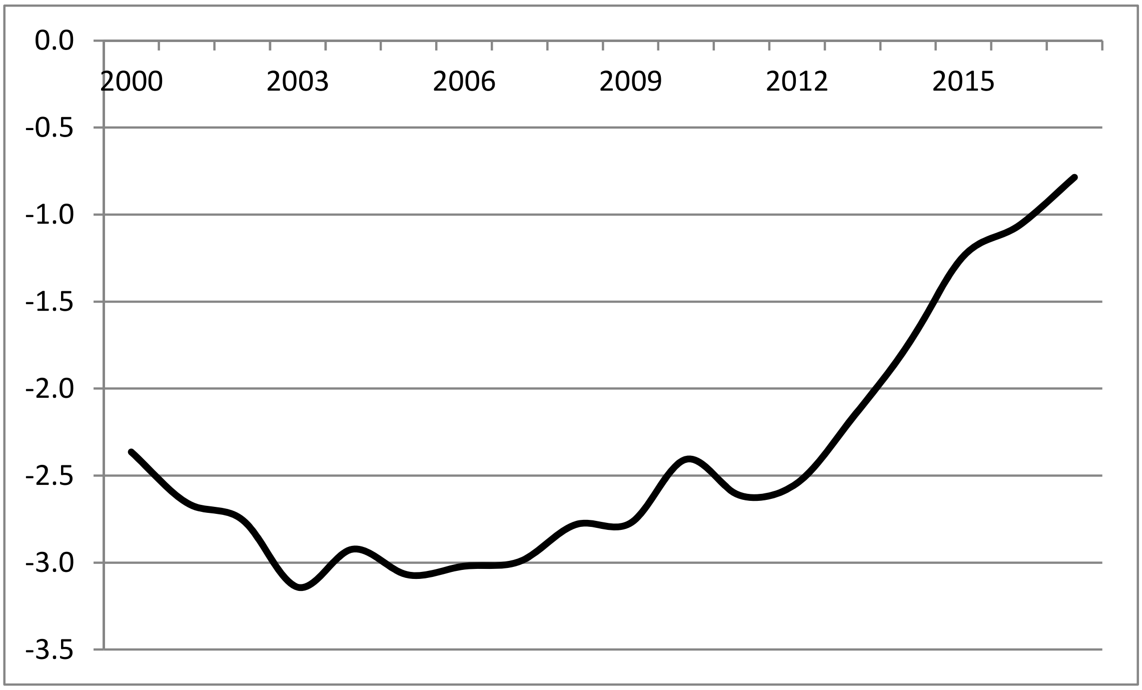
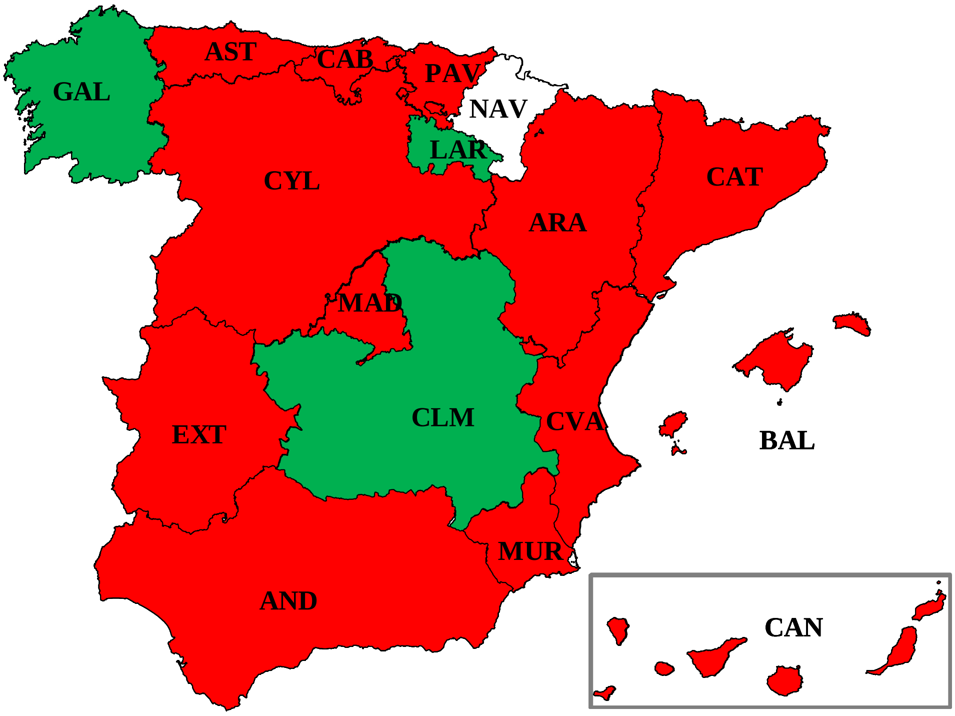


Figure 3.2. Evolution of the estimation of the parameter β.




This figure reflects the evolution of estimation of the parameter β in equation (3.3) when the sample covers the period from 1975 to N, with N= 2000, 2001, …, 2017.

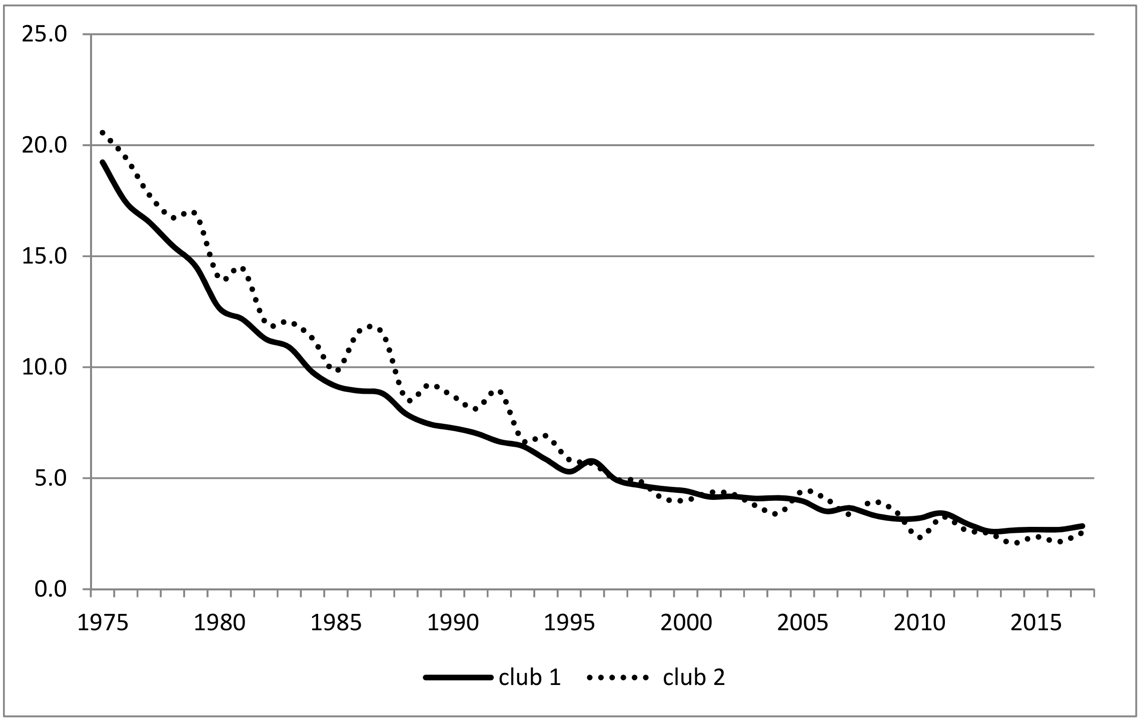
Figure 3.3. Estimated convergence clubs




This figure presents the estimated clubs.

Club 1. Club 2

Figure 3.4. Average values of the total IMR in convergence clubs 1-2



# Conclusiones

En la presente tesis hemos analizado diversos aspectos de la salud para diferentes zonzas geográficas con el común denominador de emplear siempre la misma técnica econométrica: el análisis de convergencia. La razón para su uso es que es una herramienta muy poco aplicada y que, en nuestra opinión, ofrece resultados interesantes cuando se trata de conocer la evolución temporal de diversas variables, medidas siempre como series temporales. Además, la posibilidad de agrupar las diferentes regiones de acuerdo con la evolución de estas variables de salud nos permite luego estudiar la influencia de diversas variables socioeconómicas sobre las mismas. Este aspecto es importante para, por ejemplo, poder determinar si la reciente crisis ha podido influir sobre el nivel de salud de una determinada población.

Si bien en cada capítulo se han incluido las diferentes conclusiones a las que hemos llegado, la normativa para obtener la mención internacional exige que se incluyan unas conclusiones en castellano. Esta es la razón de incluir aquí un resumen de los resultados previamente obtenidos.

El capítulo 1 de la tesis se analiza la evolución de la epidemia de VIH / SIDA en África Subsahariana. Nos hemos centrado en dos medidas diferentes de este problema: la tasa de prevalencia del VIH y la tasa de incidencia del VIH. Dada la disponibilidad de datos, hemos desglosado la tasa de prevalencia del VIH por género. El uso de los avances recientes en el análisis de convergencia nos ha permitido estudiar si estas dos medidas convergen y exhiben un comportamiento único en todo el tamaño de la muestra, que abarca 1990-2016, o si, por el contrario, exhiben diferentes patrones de comportamiento. Si nos encontramos en este último caso, entonces deberíamos concluir que la epidemia no está controlada y que algunos países aún exhiben un patrón de comportamiento diferente al resto, a pesar del gran esfuerzo realizado por los países subsaharianos con el apoyo extraordinario de comunidad internacional.

Nuestros resultados nos llevan a concluir que la prevalencia masculina del VIH y las tasas totales de incidencia del VIH muestran un patrón de comportamiento único. Esto implica que aquellos países con las tasas más altas al comienzo de la muestra han reducido la distancia con respecto a aquellos con las tasas más pequeñas al disminuir sus tasas de prevalencia / incidencia, lo que podría interpretarse como el éxito de las políticas adoptadas. Además, hemos verificado que este resultado es relativamente reciente, en el sentido de que se saca una conclusión diferente cuando se considera la muestra 1990-2010, al menos para la tasa de incidencia total de VIH. En consecuencia, parece que todas las políticas adoptadas, que fueron definidas por la agenda de los ODM, han sido bastante efectivas en la lucha contra la epidemia de VIH / SIDA y ambas tasas; MPHIV y TIHIV, parecen estar bajo control.

Sin embargo, podemos rechazar la presencia de un patrón de comportamiento único para la tasa de prevalencia del VIH femenino. Por el contrario, observamos que la tasa femenina de los países con la tasa de prevalencia más alta (Botswana, Lesotho, Malawi, Mozambique, Namibia, Sudáfrica, Swazilandia, Uganda, Zimbabwe y Zambia) más Guinea Ecuatorial y Sierra Leona, con un crecimiento comparativamente elevado en los últimos años, se comporta de manera diferente con respecto al resto de los países subsaharianos. Por lo tanto, este análisis de convergencia nos ayuda a identificar la existencia de un grupo de países donde el VIH / SIDA no parece estar controlado todavía.

Habiendo encontrado el origen de las diferencias, tratamos de determinar qué factores pueden conducirlas. Con ese fin, hemos estimado un modelo probit con la variable dependiente tomando el valor 0 si el país pertenece al club 2 (el club más grande) y 1 si el país pertenece al club 1 (el club con los países que exhiben el mayor VIH femenino tasas de prevalencia). Hemos utilizado varias variables explicativas que capturan los diferentes aspectos socioeconómicos de estos países, principalmente relacionados con la educación, la economía y los hábitos sexuales, todos ellos comúnmente identificados como factores clave para comprender la transmisión del VIH / SIDA. Las variables finalmente incluidas en el modelo son el porcentaje de población urbana, la tasa de desempleo juvenil masculino, el porcentaje de la población femenina de 15-24 años que usó un condón en su última relación sexual en los últimos 12 meses y el porcentaje de niñas inscritas en el primer año de primaria que eventualmente llegan al último año de educación primaria. Los resultados obtenidos implican que la probabilidad de pasar del club 1 al club 2 aumenta con la mejora de la situación económica (como lo indica el la tasa de desempleo), el uso más consistente de los métodos anticonceptivos y, especialmente, el incremento del nivel educativo femenino. Hemos analizado la evolución de la probabilidad de ser incluido en el club 1 cuando cambian las variables socioeconómicas. La combinación de un aumento de la persistencia femenina al último año de primaria y una reducción de la tasa de desempleo juvenil masculino a su media muestral respectiva sería bastante efectiva y reduciría la probabilidad de ser asignada al club 1 por debajo del 50%. Este resultado refuerza la idea de que las mejoras en la educación general, especialmente aquellas dirigidas a la población femenina, son cruciales para erradicar la pandemia del VIH / SIDA, aunque deberían ir acompañadas de la toma de decisiones apropiadas sobre las políticas económicas y de salud.

Finalmente, debemos tener en cuenta que la adopción de nuestras políticas sugeridas debe ir acompañada de la continuación del apoyo de la comunidad internacional. Debemos reconocer que el progreso logrado en la lucha contra el VIH / SIDA se ha logrado parcialmente debido a una inversión financiera muy grande de la comunidad internacional, como señalan Oberth y Whiteside (2016). Si no se mantiene este apoyo, no parece posible que estos países proporcionen recursos suficientes para la lucha contra el VIH / SIDA y, en consecuencia, todos los logros podrían sufrir un serio revés.

En el capítulo 2 se estudia la convergencia de las tasas mundiales de mortalidad infantil. Consideramos dos medidas diferentes, la tasa de mortalidad infantil y la tasa de mortalidad de menores de 5 años. Nuestros resultados nos llevan a rechazar la existencia de un patrón de comportamiento único en ambas tasas. La evolución de estas tasas para los países considerados puede, de hecho, agruparse en hasta 10 clubes de convergencia diferentes. Los países africanos y asiáticos se incluyen principalmente en los clubes con las tasas de mortalidad más altas, mientras que los países europeos pertenecen a los clubes con las tasas más bajas, aunque las tasas de Cuba, Corea, Japón y Singapur son similares a las de los países europeos.

Al analizar por qué un país en particular está incluido en un club, los resultados de la estimación de un modelo probit ordenado muestran que esta clasificación depende de la densidad de población, la tasa de fecundidad, el PIB per cápita, el acceso a servicios básicos de saneamiento, el porcentaje de población urbana y el total tasa de desempleo. Nuestros resultados confirman la existencia de una relación inversa entre la fecundidad y la tasa de mortalidad infantil.

Finalmente, nuestras conclusiones ofrecen resultados interesantes para los responsables políticos. Hemos visto que la creación de un club depende de la situación económica y, especialmente, del acceso a los servicios básicos de saneamiento. Además, hemos analizado cómo las variaciones en las variables explicativas cambian las probabilidades estimadas de ser incluidos en los diferentes clubes y cómo estas variaciones pueden ayudarnos a predecir los valores de las tasas de mortalidad infantil. Los resultados de este documento nos llevan a concluir que es poco probable que se logren los ODS en 2030. La reducción de la TMI a menos de 12 muertes por cada 1,000 nacimientos vivos requeriría incrementos improbables del PIB per cápita. Incluso si el cumplimiento de los objetivos se pospusiera para 2050, las economías tendrían que crecer a tasas muy improbables.

Por último, el capítulo 3 analiza la evolución de la tasa de mortalidad infantil regional española. Nuestros resultados nos llevan a rechazar esta hipótesis, lo que implica que estas tasas de mortalidad no comparten un patrón de comportamiento similar. Más bien, podemos observar la presencia de dos clubes de convergencia diferentes.

También hemos considerado el posible efecto de los recortes en los servicios de salud derivados de la Gran Recesión en la evolución de las tasas de mortalidad infantiles. Si bien es cierto que la hipótesis nula de convergencia se rechaza incluso cuando la muestra solo considera el período anterior a la Gran Recesión, podemos observar que el incremento de la muestra reduce la evidencia contra la hipótesis nula. Por lo tanto, parece haber una conexión entre las decisiones económicas y la evolución de la tasa de mortalidad infantil. Además, si la tendencia actual continúa, podríamos esperar convergencia en el futuro cercano. Pero, este proceso de convergencia iría en la dirección opuesta, dado que las regiones con las tasas de mortalidad infantiles más bajas aumentarían sus tasas, alcanzando a las regiones con las tasas más altas.

Finalmente, hemos estimado un modelo para saber qué fuerzas son las que impulsan la creación de los clubes de convergencia. Usando un probit ordenado, observamos que la probabilidad de que una región sea asignada al club 1, la que tiene la tasa de mortalidad infantil más elevada, crece con el número de mujeres que admiten haber bebido bebidas alcohólicas en las últimas dos semanas, la densidad de población y tasa de desempleo total. La inclusión de esta última variable podría sugerir nuevamente que los recortes en los servicios de salud y la evolución de las tasas de mortalidad infantil están relacionados. Sin embargo, este último resultado debe interpretarse con cierta cautela y se deben realizar análisis adicionales, especialmente un análisis de causalidad entre estas dos variables.

# Appendix A.

## A1. List of the variables and data source

| WHO Code | Indicator Name | Sample |
|---|---|---|
| EN.URB.MCTY.TL.ZS | Population in urban agglomerations of more than 1 million (% of total population) | Average of the 2000-2016 period. |
| NY.ADJ.NNTY.PC.KD | Adjusted net national income per capita (constant 2010 US$) | Average of the 2000-2016 period |
| NY.GDP.PCAP.PP.KD | GDP per capita, PPP (constant 2011 international $) | Average of the 2000-2016 period |
| SE.PRM.PRSL.FE.ZS | Persistence to last grade of primary, female (% of cohort) | Average of the 2000-2016 period |
| SE.PRM.PRSL.MA.ZS | Persistence to last grade of primary, male (% of cohort) | Average of the 2000-2016 period |
| SE.PRM.PRSL.ZS | Persistence to last grade of primary, total (% of cohort) | Average of the 2000-2016 period |
| SE.SEC.PROG.FE.ZS | Progression to secondary school, female (%) | Average of the 2000-2016 period |
| SE.SEC.PROG.MA.ZS | Progression to secondary school, male (%) | Average of the 2000-2016 period |
| SE.SEC.PROG.ZS | Progression to secondary school (%) | Average of the 2000-2016 period |
| SH.CON.1524.FE.ZS | Condom use, population ages 15-24, female (% of females ages 15-24) | Last observed value in the 1990-2016 sample |
| SH.CON.1524.MA.ZS | Condom use, population ages 15-24, male (% of males ages 15-24) | Last observed value in the 1990-2016 sample |
| SH.XPD.PCAP.PP.KD | Health expenditure per capita, PPP (constant 2011 international $) | Average of the 2000-2016 period |
| SH.XPD.PUBL.ZS | Health expenditure, public (% of GDP) | Average of the 2000-2016 period |
| SH.XPD.TOTL.ZS | Health expenditure, total (% of GDP) | Average of the 2000-2016 period |
| SI.POV.GINI | GINI index (World Bank estimate) | Average of the 2000-2016 period |
| SL.UEM.1524.FE.ZS | Unemployment, youth female (% of female labor force ages 15-24) (modeled ILO estimate) | Average of the 2000-2016 period |
| SL.UEM.1524.MA.ZS | Unemployment, youth male (% of male labor force ages 15-24) (modeled ILO estimate) | Average of the 2000-2016 period |

This table includes the list of the explanatory variables that we have employed to estimate the ordered probit model. All of them have been obtained from the World Development Indicators Database, with the sample covering 2000-2016.

The countries included are those reported in Table 1.1, except Angola, Botswana, Cabo Verde, Central Africa Republic, Equatorial Guinea, Eritrea, Gabon, Guinea Bissau Mauritania, Somalia, South Sudan and Sudan, due to lack of data.

# Appendix B

## B1. Source of the data.

| Variable | Indicator Name | Source |
|---|---|---|
| Battle-related deaths | Battle-related deaths (number of people) | Uppsala Conflict Data Program, http://www.pcr.uu.se/research/ucdp/. |
| Presence of peacekeepers | Presence of peace keepers (number of troops, police, and military observers in mandate) | UN Department of Peacekeeping Operations, http://www.un.org/en/peacekeeping/. |
| Intentional Homicides | Intentional homicides (per 100,000 people) | UN Office on Drugs and Crime's International Homicide Statistics database. |
| Population Density | Population density (people per sq. km of land area) | Food and Agriculture Organization and World Bank population estimates. |
| Urban Population | Urban population (% of total) | The United Nations Population Division's World Urbanization Prospects. |
| Births | Births attended by skilled health staff (% of total) | UNICEF, Country of the World's Children, Childinfo, and Demographic and Health Surveys. |
| Fertility rate | Fertility rate, total (births per woman) | United Nations Population Division. World Population Prospects. |
| Improved sanitary facilities | Improved sanitary facilities (% of population with access) | WHO/UNICEF Joint Monitoring Programme (JMP) for Water Supply and Sanitation |
| Per capita GDP | GDP per capita (constant 2010 US$) | World Bank national accounts data, and OECD National Accounts data files. |
| Industry (% GDP) | Industry, value added (% of GDP) | World Bank national accounts data, and OECD National Accounts data files. |
| Female Unemployment rate | Unemployment, female (% of female labor force) | International Labour Organization, ILOSTAT database |
| Male Unemployment rate | Unemployment, male (% of male labor force) | International Labour Organization, ILOSTAT database. |
| Total Unemployment rate | Unemployment, total (% of total labor force) | International Labour Organization, ILOSTAT database. |
| Private health expenditure (%) | Health expenditure, private (% of GDP) | World Health Organization Global Health Expenditure database |
| Public health expenditure (%) | Health expenditure, public (% of government expenditure) | World Health Organization Global Health Expenditure database |
| Total health expenditure (%) | Health expenditure, total (% of GDP) | World Health Organization Global Health Expenditure database |
| Female literacy rate | Literacy rate, adult female (% of females ages 15 and above) | United Nations Educational, Scientific, and Cultural Organization (UNESCO) |
| Infant mortality rate | Mortality rate, infant (per 1,000 live births) | Estimates Developed by the UN Inter-agency Group for Child Mortality Estimation |
| Under.5 mortality rate | Mortality rate, under-5 (per 1,000 live births) | Estimates Developed by the UN Inter-agency Group for Child Mortality Estimation |

**B2. Ordered probit estimation**

An ordered probit model has been used to analyze the formation of the different clubs in which the dependent variable $y_i$ is defined as follows:

$$y_i = m \text{ for } m=1,...,M \tag{1}$$

with M being equal to 9 and 10 for IMR and U5MR, respectively. The observed values are assumed to derive from some unobservable latent variable $y_i^*$ where:

$$y_i^* = x_i'\beta + u_i, \quad i=1,2,\ldots,154 \tag{2}$$

for some $k\times 1$ parameter vector $\beta$, a vector of explanatory variables ($x_i$) of dimension k and univariate stochastic disturbance term $u_i$.

We interpret that the different $M$ values imply a preference or a ranking of the clubs and, therefore, the observed variable, $y_i$, is related to the latent variable through the following observability criterion:

$$y_i = m, \text{ if } \alpha_{m-1} \le y_i^* \le \alpha_m, \; j=1, 2, \ldots M \tag{3}$$

for a set of parameters $\alpha_0$ to $\alpha_M$ , where $\alpha_o<\alpha_1<\alpha_2...<\alpha_M$, with $\alpha_o=-\infty$ and $\alpha_M=\infty$.

Then, the conditional probability of observing the m-th category can be written as:

$$Pr_m=Pr(y_i=m/x_i)=Pr(\alpha_{m-1}\le y_i^*\le\alpha_m) = Pr(\alpha_{m-1}\le x_i'\beta + u_i \le\alpha_m) =$$

$$= Pr(u_i\le\alpha_m - x_i'\beta) - Pr(u_i \le \alpha_{m-1} - x_i'\beta). \tag{4}$$

To evaluate the conditional probability, we assume that the perturbation follows a normal distribution.

Thus, we can estimate the probability of any country being in club m by simply taking into account the estimated values presented in Table 2.5. Moreover, we can easily predict the value of the two child mortality rates as follows:

$$\widehat{CMR} = \sum_{m=1}^{M} \widehat{Pr_m}\ \overline{CMR_m} \qquad (5)$$

with $\widehat{Pr_m}$substituting the parameters β and α by their corresponding estimations and $\overline{CMR_m}$being the average value of the child mortality rates for the countries included in the m-th club, where CMR= {IMR, U5MR}.

In order to evaluate this method of prediction, we can analyze the following case. Let us create 5 dummy countries which represent the average behavior of the countries in the five continents. If we follow this procedure, and bearing in mind that the average values of the IMR for each club are { 61.2, 37.5, 19.1, 10.9, 6.3, 3.3,,2.8, 2.0,1.6}, then the predicted IMR for 2015 are 43.2, 15.1, 15.0, 21.0 and 4.2 for the average-type African, American, Asian, Oceanian and European country, respectively. The actual observed values are 45.7, 15.8, 19.7, 22.8 and 4.8, respectively, so we should conclude that the prediction is accurate.

Finally, Section 2.4 analyzes the behavior of the countries in club 1 when the explanatory variables change. In order to control for the evolution of these variables, we have considered that the explanatory variables evolve according to the following equations:

- PDEN* = 1.01x PDEN
- URB* = 1.01 x URB
- pcGDP* = (1+g) x pcGDP

- $PUBH^* = 8.7 + 0.0003 \times pcGDP^*$
- $TFR^* = 5.3 - 0.0003 \times pcGDP^*$
- $BSS^* = 20.7 + 0.007 \times pcGDP^*$
- $TUR^* = 7.4 - 0.000201 \times pcGDP^*$

with PUBH, TFR, BSS and TUR being the public health expenditure, the total fertility rate, the basic sanitation services, and the total unemployment rate, respectively. pcGDP, PDEN and URB are the average value of the per capita GDP, the population density and the urban population percentage of the countries in club 1 at the end of the sample. The slope of these functions has been obtained by regressing these variables on the per capita GDP for the countries in club 1. The intercept is the solution of the previous equations when the variables take the average value of the countries in club 1 for 2015 and g=0.

# Appendix C

## C3. Definition of the explanatory variables.

| Variable | Definition | Source |
|---|---|---|
| Terrain Ruggedness | Terrain Ruggedness Index | Fundación BBVA |
| Population density | Surface/population. 1975-2015 average. | Instituto Nacional de Estadística |
| GDP (pc) | Gross domestic product divided by population. 1980-2015 average. | Instituto Nacional de Estadística |
| Unemployment rate | 2006-2016 average | Instituto Nacional de Estadística |
| Migration rate | Flow of immigration from abroad of a region divided by the total population of the region. 2008-2016 average | Instituto Nacional de Estadística |
| Public Health Exp. (pc) | Total public health expenditure divided by population. Average 1991-2010. | Fundación BBVA |
| Private medical assurance rate | Percentage of the population exclusively having a private medical assurance. | Instituto Nacional de Estadística. Encuesta Nacional de Salud. Año 2011-2012 |
| Female Obesity rate | Percentage of the female population older than 15 with a body mass index greater (or equal) than 30 kg/m2 | Instituto Nacional de Estadística. Encuesta Nacional de Salud. Año 2011-2012 |
| Female Tobacco rate | Percentage of the female population older than 15 that admits smoking at least a cigarette a day. | Instituto Nacional de Estadística. Encuesta Nacional de Salud. Año 2011-2012 |
| Female alcoholic beverage consumption rate | percentage of the female population older than 15 that admits having drunk alcoholic beverages in the last two weeks | Instituto Nacional de Estadística. Encuesta Nacional de Salud. Año 2011-2012 |
| Fertility global rate | Total number of births by 1,000 women in fertile age (15-49). Average of the 1975-2015 period. | Instituto Nacional de Estadística. |

# References

Alsan, M. and D.M. Cutler. (2013). Girls' education and HIV risk: evidence from Uganda. Journal of health economics, vol. 32, no 5, p. 863-872.

Alves, D., & Belluzzo, W. (2004). Demand for child health in Brazil: a household decision model. Final seminar on child health, poverty and the role of social policies. Washington, DC: IDB.

Amaugo, L. G., Papadopoulos, C., Ochieng, B. M., & Ali, N. (2014). The effectiveness of HIV/AIDS school-based sexual health education programmes in Nigeria: A systematic review. Health education research, 29(4), 633-648.

Asiedu, C., Asiedu, E., & Owusu, F. (2012). The Socio-Economic Determinants of HIV/AIDS Infection Rates in Lesotho, Malawi, Swaziland and Zimbabwe. Development Policy Review, 30(3), 305-326.

Barro R. and X. Sala-i-Martin, (1991). Convergence across countrys and regions. Brookings Papers on Economic Activity, 1, 107-158.

Barro, R. and Sala-i-Martin, X. (1992). Convergence. Journal of Political Economy, 100, 223-251.

Baumol, W. J. (1986). Productivity Growth, Convergence, and Welfare: What the Long-Run Data Show. The American Economic Review, Vol. 76, No. 5 (Dec., 1986), pp. 1072-1085

Behrman, J. A. (2015). The effect of increased primary schooling on adult women's HIV status in Malawi and Uganda: Universal Primary Education as a natural experiment. Social science & medicine, 127, 108-115.

Bremberg, S. G. (2017). Mortality rates in OECD countries converged during the period 1990–2010. Scandinavian Journal of Public Health, 1403494816685529.

Bunyasi, E. W., & Coetzee, D. J. (2017). Relationship between socioeconomic status and HIV infection: findings from a survey in the Free Country and Western Cape Provinces of South Africa. BMJ open, 7(11), e016232.

Chen F, Bajwa NM, Rimensberger PC and the Swiss Neonatal Network, et al. (2016). Thirteen-year mortality and morbidity in preterm infants in Switzerland. Archives of Disease in Childhood - Fetal and Neonatal Edition, 101(5), F377-F383.

De Long, J. B (1988). Productivity Growth, Convergence, and Welfare: Comment. The American Economic Review, Vol. 78, No. 5 (Dec., 1988), pp. 1138-1154.

De Neve, J. W., Fink, G., Subramanian, S. V., Moyo, S., & Bor, J. (2015). Length of secondary schooling and risk of HIV infection in Botswana: evidence from a natural experiment. The Lancet Global Health, 3(8), e470-e477.

Defo, B. K. (1996). Areal and socioeconomic differentials in infant and child mortality in Cameroon. Social Science and Medicine, 42(3), 399–420.

Dyson, T. (2003). HIV/AIDS and urbanization. Population and Development Review, 29(3), 427-442.

Fink, G., Günther, I., & Hill, K. (2011). The effect of water and sanitation on child health: Evidence from the Demographic and Health Surveys 1986–2007. International Journal of Epidemiology, 40, 1196–1204.

Fink, G., Günther, I., & Hill, K. (2014). Slum residence and child health in developing countries. Demography, 51, 1175–1197.

Gbesemete, K. P., & Jonsson, D. (1993). A comparison of empirical models on determinants of infant mortality: a cross-national study on Africa. Health Policy, 24, 155–174.

Gil-Alana, L. A., Cunado, J., & Gupta, R. (2017). Persistence, mean-reversion and non-linearities in infant mortality rates. Social Indicators Research, 131, 393-395.

Gil-Alaña, L. A., Cunado, J., & Gupta, R. (2017). Persistence, mean-reversion and non-linearities in infant mortality rates. Social Indicators Research, 131, 393-395.

Glynn, J.R., Carae, M., Buve, A., Anagonou, S., Zekeng, L., Kahindo, M., Musonda, R., (2004). Does increased general schooling protect against HIV infection? A study in four African cities. Trop. Med. Int. Health. 9 (1), 4e14.

Grogan, L., (2008). Universal primary education and school entry in Uganda. J. Afr. Econ. 18 (2), 183e211.

Hajizadeh, M., Sia, D., Heymann, S. J., & Nandi, A. (2014). Socioeconomic inequalities in HIV/AIDS prevalence in sub-Saharan African countries: evidence from the Demographic Health Surveys. International journal for equity in health, 13(1), 18.

Hakobyan, M., Mkrtchyan, A., & Yepiskoposyan, L. (2006). Infant mortality in Armenia, 1992-2003. Economics and Human Biology, 4, 351–358.

Hanmer, L., Lensink, R., & White, H. (2003). Infant and child mortality in developing countries: analyzing the data for robust determinants. The Journal of Development Studies, 40(1), 101–118.

Hargreaves, J.R., Bonell, C.P., Boler, T., et al., (2008). Systematic review exploring time trends in the association between educational attainment and risk of HIV infection in sub-Saharan Africa. AIDS 22, 403e414.

Hathi, P., Haque, S., Pant, L., Coffey, D., & Spears, D. (2014). Place and child health: the interaction of population density and sanitation in developing countries. Demography, 1, 1-24.

Hondroyiannis, G., & Papapetrou, E. (2002). Demographic transition in Europe. Economics Bulletin, 10(3), 1–8.

Jankowska, M. M., Benza, M., & Weeks, J. R. (2013). Estimating spatial inequalities of urban child mortality. Demographic Research, 28(article 2), 33–62. doi:10.4054/DemRes.2013.28.2

Jukes, M., Simmons, S., Bundy, D. (2008). Education and vulnerability: the role of schools in protecting young women and girls from HIV in southern Africa. AIDS (Suppl. 4), S41eS56.

Kimani-Murage, E. W., Fotso, J. C., Egondi, T., Abuya, B., Elungata, P., Ziraba, A. K. & Madise, N. (2014). Trends in childhood mortality in Kenya: the urban advantage has seemingly been wiped out. Health & place, 29, 95-103.

Ko, Y. J., Shin, S. H., Park, S. M., Kim, H. S., Lee, J. Y., Kim, K. H., et al. (2014). Effects of employment and education on preterm and full-term infant mortality in Korea. Public Health, 128(3), 254–261.

Lantto, M. (2015). Childhood mortality in Finland. Väitöskirja. Lääketieteenlaitos. Oulun yliopisto. Viitattu, 7, 2016.

Leclerc-Madlala, S., Simbayi, L. C., & Cloete, A. (2009). The sociocultural aspects of HIV/AIDS in South Africa. In HIV/AIDS in South Africa 25 Years On (pp. 13-25). Springer, New York, NY.

Leon, J., Baker, D. P., Salinas, D., & Henck, A. (2017). Is education a risk factor or social vaccine against HIV/AIDS in sub-Saharan Africa? The effect of schooling across public health periods. Journal of Population Research, 34(4), 347-372.

Mabaso, M., Sokhela, Z., Mohlabane, N., Chibi, B., Zuma, K., & Simbayi, L. (2018). Determinants of HIV infection among adolescent girls and young women aged 15–24 years in South Africa: a 2012 population-based national household survey. BMC public health, 18(1), 183.

Magadi, M. A., & Magadi, W. A. (2017). HIV/AIDS and contraceptive use: factors associated with contraceptive use among sexually active HIV-positive women in Kenya. Contraception, 95(3), 312-321.

Marlow, N. (2014). Interpreting regional differences in neonatal outcomes for extremely preterm babies. Acta Paediatrica, 103(1), 4-5.

Matsaganis, M. (1992). Maternal mobility and infant mortality in Greece: A regional analysis. Social science & medicine, 34(3), 317-323.

Mbirimtengerenji ND. (2007). Is HIV/AIDS Epidemic Outcome Of poverty in sub-Saharan Africa? Croat Med J., 48(5):605–17.

McMichael AJ, McKee M, Shkolnikov V, Valkonen T. (2004). Mortality trends and setbacks: global convergence or divergence? Lancet, 363:1155-9.

Mondal, N. I., Hossain, K., & Ali, K. (2009). Factors influencing infant and child mortality: A case study of Rajshahi District, Bangladesh. Journal of Human Ecology, 26(1), 31–39.

Moser, K., Shkolnikov, V., & Leon, D. A. (2005). World mortality 1950-2000: divergence replaces convergence from the late 1980s. Bulletin of the World Health Organization, 83(3), 202-209.

Mufune P. (2014). Poverty and HIV/AIDS in Africa: specifying the connections. Social Theory & Health, 13(1):1–29.

Muldoon, K. A., Galway, L. P., Nakajima, M., Kanters, S., Hogg, R. S., Bendavid, E. & Mills, E. J. (2011). Health system determinants of infant, child and maternal mortality: a cross-sectional study of UN member countries. Globalization and health, 7(1), 42.

Narayan, P. K., & Peng, X. (2007). Japan's fertility transition: empirical evidence from the bounds testing approach to cointegration. Japan and the World Economy, 19, 263–278.

Oberth, G., & Whiteside, A. (2016). What does sustainability mean in the HIV and AIDS response? African Journal of AIDS Research, 15(1), 35-43.

Panopoulou, E. and Pantelidis, T. (2013). Cross-country disparities in us health care expenditures. Health Economics, 22: 451-465.

Pettifor, A.E., Rees, H.V., Kleinschmidt, I., et al. (2005). Young people's sexual health in South Africa: HIV prevalence and sexual behaviors from a nationally representative household survey. AIDS 19, 1525e1534.

Phillips P C B, Sul D. (2007). Transition modeling and econometric convergence tests Econometrica 756: 1771-1855.

Phillips P C B, Sul D. (2009). Economic transition and growth. Journal of Applied Econometrics 247: 1153–1185.

Quah, D. (1993). Empirical cross-section dynamics in economic growth. European Economic Review, 37 (1993) 426434. North-Holland

Quansah, E., Ohene, L. A., Norman, L., Mireku, M. O., & Karikari, T. K. (2016). Social factors influencing child health in Ghana. PloS one, 11(1), e0145401.

Regidor E., Gutierrez-Fisac J. L. and Rodriguez C. (1994). Diferencias y desigualdades en salud en España, p. 145. Diaz de Santos, S. A., Madrid, 1994.

Regidor, E., Calle, M.E., Domínguez, V., Navarro, P. (2002). Inequalities in mortality in shrinking and growing areas. Journal of Epidemiology and Community Health, 56, 919–921.

Regidor, E., Gutierrez-Fisac, J.L., Rodriguez, C. (1995). Increased socioeconomic differences in mortality in eight Spanish provinces. Social Science & Medicine. 41,801–807.

Renton, A., Wall, M., & Lintott, J. (2012). Economic growth and decline in mortality in developing countries: an analysis of the World Bank development datasets. Public Health, 126(7), 551–560.

Sastry, N. (1997). What explains rural-urban differentials in child mortality in Brazil?. Social science & medicine, 44(7), 989-1002.

Schell, C. O., Reilly, M., Rosling, H., Peterson, S., & Mia Ekström, A. (2007). Socioeconomic determinants of infant mortality: a worldwide study of 152 low-, middle-, and high-income countries. Scandinavian journal of public health, 35(3), 288-297.

Serenius, F. , Sjörs, G. , Blennow, M. , Fellman, V. , Holmström, G. , Maršál, K. , Lindberg, E. , Olhager, E. , Stigson, L. , Westgren, M. , Källen, K. and , (2014), EXPRESS study shows significant regional differences in 1-year outcome of extremely preterm infants in Sweden. Acta Paediatrica, 103(1): 27-37.

Sidebotham, P., Fraser, J., Fleming, P., Ward-Platt, M., & Hain, R. (2014). Patterns of child death in England and Wales. The Lancet, 384(9946), 904-914.

Steinert, J. I., Cluver, L., Melendez-Torres, G. J., & Herrero Romero, R. (2017). Relationships between poverty and AIDS Illness in South Africa: an investigation of urban and rural households in KwaZulu-Natal. Global public health, 12(9), 1183-1199.

Subramaniam, T., Loganathan, N., Yerushalmi, E., Devadason, E. S., & Majid, M. (2018). Determinants of Infant Mortality in Older ASEAN Economies. Social Indicators Research, 136; 397-415.

Taaffe, J., Cheikh, N., & Wilson, D. (2016). The use of cash transfers for HIV prevention—are we there yet? African Journal of AIDS Research, 15(1), 17-25.

UNDP (2005) Human Development Report. New York: United Nations Development Programme.

Vandemoortele, J., & Delamonica, E. (2000). The ‘education vaccine’against HIV. Current issues in comparative education, 3(1), 6-13.

White, H. L., (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica 48: 817–838.

Wilson C. (2001). On the scale of global demographic convergence 1950-2000. Population and Development Review, 27:155-71.

Wilson, C. (2011). Understanding global demographic convergence since 1950. Population and Development Review, 37(2), 375-388.