HIV/AIDS Knowledge Differential by Geopolitical, Social and Economic Status: Evidence from Surveyed Children in South East Asia.

Rosalia Vazquez-Alvarez [a]

University of St.Gallen, Swiss Institute for International Economics and Applied Economic Research (SIAW)

This version: March, 2004[*]

Abstract

The HIV/AIDS epidemic is a major public health threat, with evidence showing that information campaigns are effective policy tools to control its evolution. Many of these are designed to increase children's knowledge on HIV before they become sexually active or start to experiment with intravenous drugs, with the effectiveness of such campaigns subject to the extent to which they reach all sectors in the population. This paper applies matching methods and techniques for programme evaluation to detect differentials in children's knowledge on HIV/AIDS by geopolitical, social and economic status, allowing for health risk-taking behaviour. Drawing from a sample of surveyed children in South East Asia and the Pacific, the results show support for Amartya Sen's (2002) suggested negative effect of non-democracy on HIV/AIDS awareness, although children in democracies with low development and/or low HIV-prevalence seem to be less well equipped to deal with the epidemic. Within country results suggest that gender, schooling status and living in a urban setting are determinants of knowledge differentials, while parental income or occupation are not. This study has policy implications with respect to what sectors in the population should Information Campaigns target.

JEL – Classification: C13, C14, I10, I12,I18, I28, I38,012, P51.

Keywords: HIV Prevalence, Geopolitical and Socio-Economic Evaluation, Matching on the Propensity Score, HIV/AIDS Knowledge, Health-Habit formation, Preventive Information Campaigns, East Asia and The Pacific.

---

[a] Acknowledgements: I am very grateful to UNICEF's regional director Mehr Khan, and Mark Thomas (UNICEF South East Asia Communication Officer) for allowing me to use the data unconditionally for scientific non-lucrative purposes. I am also grateful to Wirote Ngamsukkasamesri (Thailand Research International Consulting) for advice on decoding the raw data. Any errors or omissions are entirely my responsibility. Views expressed in this paper do not necessarily reflect those of UNICEF or UNICEF East Asia and Pacific Regional Office.

* Rosalia Vazquez-Alvarez, Rosalia.Vazquez-Alvarez@unisg.ch, www.siaw.unisg.ch. Address for correspondence: SIAW, University of St.Gallen, Dufourstrasse 48, St. Gallen, 9000, Switzerland.

1 Introduction

In 2003 the global HIV/AIDS epidemic killed more than 3 million people, while an estimated 5 million acquired the Human Immunodeficiency Virus (HIV), the agent that causes Acquired Immunodeficiency Syndrome (AIDS). Today, very few would question that prevention remains the pillar of any health policy that aims at stopping the spread of HIV/AIDS.[1] Preventive measures are based on information campaigns aimed at endowing the population with knowledge on the social dynamics of HIV/AIDS, the assumption being that knowledge modifies individual's behaviour in high –risk situations with respect to HIV infection. However, in any given country, information campaigns are effective at controlling the rates of infection only if they reach all social and economic sectors in the population, otherwise, the behaviour of those who engage on HIV-related risk habits would end up dominating the evolution of the epidemic.[2] The aim of this paper is to find out if geopolitical, social and economic conditions explain differentials in the probability to avoid HIV infection, where such probability is assumed to be a function of direct knowledge on HIV/AIDS and risky health-related behaviour.

Differentials in HIV-prevalence between countries can be up to 25,000%. For example, in the UK, where HIV-prevalence is representative of many Western economies, the estimated number of HIV carriers in the adult population is of 1 in every 1000 (i.e., a HIV-prevalence of 0.1%), compared to Zimbabwe where estimates suggest a HIV-prevalence of 25%. The positive effect of Western economies at informing its population, and the negative effect of a slow response in most African countries, has made governments around the world place information campaigns at the full front of their health policy agenda. This is specially

[1] Exceptions to this are the Vatican announcing through Cardinal A.L. Trujillo that scientist could show how HIV was permeable to condom, these having tiny hold through which the virus could pass (BBC1 Panorama, October 2003), reinforcing the Vatican's opposition to the use of contraceptives. Another example where prevention was undermined until recently is South Africa, with its president Mbeki publicly questioning the scientific evidence that linked the virus to AIDS, a denial that has lead to one of the highest HIV infections rates in the World.

[2] See, for example, UNAIDS (2003) reporting on Thailand's increase in HIV positive cases among wives of men that have modified their behaviour (increase condom use and reduced number of visits) with sex workers, but not with their spouses. Thus, targeting only commercial sex workers and their clients may fail the aims of the campaign.

true in regions of the world where the epidemic emerged with a time-lag relative to the Western world and Africa. One such region is South East Asia and the Pacific, where scientist hope that information campaigns will have a positive impact at controlling a HIV/AIDS epidemic that already accounts for some 20% of the total HIV-positive population in the world. South East Asia and the Pacific is of particular worry to scientist monitoring the epidemic because it holds one quarter of the world's population, with 50% below the age of 20, and, therefore, more likely to engage on HIV risk-related situations (e.g., risky sexual behaviour or the use of intravenous drugs). Furthermore, poverty and socio-economic inequality is rampant in this part of the world, both within country and between neighbouring economies. While regional economic differences could lead to differentials in HIV-prevalence between countries, within-country inequality determines differentials in HIV prevalence by socio-economic status, with poverty and inequality further exacerbating the chances to avoid infection for the very poor. If information campaigns are effective at promoting knowledge on how to avoid HIV infection, these can be seen as effective policy tools to reduce health and economic inequalities in the long run, both within and between countries in the region. This is specially true if such campaigns equip children with the correct knowledge on how to avoid infection before they become sexually active or start to experiment with intravenous drugs, but also if such information reaches all children of all socio-economic backgrounds.

This paper analyses children's responses to a survey conducted by UNICEF in the first half of 2001.The survey covers 14 countries and 3 territories of South East Asia and the Pacific. Surveyed children range from the age of 9 to 17 at the time of the survey, with those 14 and older answering questions that provide an error free indicator on the child's level of knowledge with respect to the different channels of HIV/AIDS transmission. The survey is also indicative on each child's probability to be engaged on health related risky behaviour (e.g., smoking, glue-sniffing or use of illicit drugs). Allowing for knowledge of HIV/AIDS, reported health behaviour, and given a country's specific HIV-prevalence, this paper proposes a (within-group) relative measure on the probability that a child is badly equipped to avoid HIV infection, P_A , assuming that present knowledge and risk behaviour are the best indicators for present and (near-)future behaviour. The constructed measure P_A is continuous on the [0,1] interval and reflects the effectiveness of information campaigns in the region (see Section 3.2).

The survey is also informative with respect to economic indicators (parental income, occupation, possession of particular wealth items), and socio-demographic characteristics (urban/rural household, household size, child's gender and the child's going-to-school status, among others). This paper adopts the techniques of epidemiologists based on studying the effect of a treatment (e.g., belonging to a particular geopolitical, social or economic class) on the outcome of interest (i.e., the assigned (relative) probability of becoming HIV positive, P_A). These techniques are established tools for programme evaluation in social sciences when the outcome of interest is often measured with non-experimental (rather than laboratory) data.[3] The idea is to use nonparametric matching methods (e.g., matching on the propensity score) to estimate the counterfactual outcome to the treated population. The use of nonparametric tools means that results are free from model misspecification bias, while using matching methods implies a check on common support conditions, thus improving on regression analysis that could in practice project the counterfactual outcome outside the common support.

The results show that, on average, children in South East Asia and the Pacific region are poorly equipped to deal with HIV/AIDS. Overall, only 13.5% show combined knowledge and risk behaviour that would imply safe conducts in situations where HIV might pose a risk. Treatment analysis suggest that, relative to the rest of the region, children in Low Income countries display knowledge and risk behaviour that makes them 7% worse off in terms of avoiding HIV infection. By a similar method, children in Western style economies are 5% worse off, relative to all other styles, whereas children in countries of Communist style are 9.3% better off than their counterparts in other countries of the region. All three figures are significant even at 1% level. To make sure that 'risk-behaviour' alone is not what might be driving the results (e.g., children in Western and Low Income countries might be more exposed to illegal drug use), a similar analysis is conducted but ignoring risk behaviour when constructing P_A and allowing only for HIV/AIDS knowledge.

The results are qualitatively similar, with children in Communist style countries showing knowledge that makes them 6.1% better off – relative to other children in the region – at keeping safe in a situation where HIV might pose a risk. One explanation for this is that too many countries in the survey performed very badly while the few that do relatively better in the Communist block drive the results. To understand the true performance of children in Communist style economies, the paper compares each country in the Communist block (independently) to Thailand, a democracy geographically surrounded by communist–like economies that has used information campaigns since its first HIV cases were diagnosed in the early 1990s. Except for both Viet Nam and Cambodia, and relative to their Thai counterparts, children in all other 4 Communist economies (China, Myanmar, Mongolia and PDR Laos) are severely badly equipped to avoid HIV infection. This result agrees with Amartya Sen's (see Sen (2002) and Manning (2002)) theoretical assessments on education as a primary factor to reduce HIV/AIDS inequality, and Sen's prediction that non-democratic regimes may not be sufficiently responsive to a HIV/AIDS epidemic.

A second set of results looks at within-country effects of various socio-economic indictors. These suggest that living in a urban (rather than rural) setting, going to school and been a girl significantly increase the relative probability of avoiding HIV infection, while household income and occupational skill of the main (adult) household earner do not have a significant impact on a child's P_A value.

The reminder of the paper is organized as follows. Section 2 explains the affect of information campaigns on the evolution of the HIV/AIDS epidemic. Section 3 presents the data used in the empirical analysis, with a detail explanation on the construction of P_A and a description of the control variables and treatment indicators. Section 4 describes the econometric methodology and identifying conditions. Section 5 presents the results of testing for P_A differentials, between and within countries, conditioning on selected socio-economic indicators and children's attitudes. Section 6 concludes.

[3] See Angrist (1991), Heckman, Ichimura and Todd (1997), Black and Smith (2003), Lechner (1999) or Lechner and Vazquez-Alvarez (2003) as examples of applied work on the econometric evaluation of social interventions.

2 Evidence of the effect of Information Campaigns on the Evolution of the HIV/AIDS Epidemic.

According to UNAIDS, the United Nations agency in charge of monitoring the HIV/AIDS epidemic, by the end of 2003 there were, worldwide, between 40 and 45 million people infected with HIV (the virus that leads to AIDS), while, since the onset of the global epidemic in 1981, some 20 million people have died due to AIDS or related health outcomes.[4] These evidence are sufficient to demonstrate that HIV/AIDS is today's greater health threat, with clear social, demographic and economic implications. However, the effect of such implications have experienced a social mutation over the past two decades. In the 1980s HIV/AIDS affected wealthy, white gay-man and was driven by Men-to-Men-Transmission (MTMT in epidemiology) in wealthy parts of the world (Western Europe and the West Coast of the USA). Today, the epidemic affects mostly countries in Sub-Saharan Africa, Asia and Latin America, regions identified by the UN as the worst hit areas in the world. These regions alone account for 98% of the global HIV-positive population, with Sub-Sahara holding at least 71% of the total, followed by East Asia and the Pacific with some 20% of the total. Furthermore, HIV transmission in wealthy nations is still mostly due to MTMT or among Intravenous Drug Users (IDUs), whereas in the developing world HIV/AIDS is mostly the result of heterosexual transmission, and, therefore, directly or indirectly affecting vulnerable groups in society such as women and children.

It is well known that the massive differentials on the evolution of the epidemic between developed and developing economies is due to the fast reaction of governments in wealthy nations at the onset of the epidemic. For example, in the UK, as early as 1985 there were nationwide information campaigns designed to reach all sectors in the population, effectively modifying people's behaviour to reduce individual's exposure to the risk of infection. Today, the UK has a very low prevalence (0.1%) and a high degree of control over the virus (e.g., authorities know that two thirds of the HIV-positive population live in the London region,

[4] Some worst case scenario predicted the total number of HIV positive in the world at 110 million. See Dow and Philipson (1996) quoting worldwide predictions from the Global AIDS Policy Coalition (1992).

knowledge that allows better predictions on the epidemic's evolution and on information targeting in the population). Similar comments apply to other countries in West Europe and the USA[5]

On the other hand, developing economies, specially Sub-Saharan countries, have seen exponential increases in their HIV-positive populations and a systematic reduction in life expectancy due to AIDS related deaths.[6] The onset of the epidemic in Africa was at a similar time and with similar HIV-prevalence than in Western Europe. The difference, with few exceptions,[7] is that health authorities in African countries did not react with the required information campaigns. One of the exceptions was Uganda where following the first diagnosed cases in the mid 1980s an energetic campaign promoting abstinence, fidelity and condom use, resulted in a drop of HIV incidence from 30% in the early onset to 5% in 2001 (De Walque (2001)). The opposite happened in South Africa. In 1988, South Africa and many other countries in Africa, showed a HIV-prevalence in adults of 2%. In South Africa, both pre-Apartheid and post-Apartheid administrations ignored

[5] The highest rates in Europe are in Switzerland and Spain, both with a HIV-prevalence of 0.5% in the adult population. Other countries fluctuate in the 0.05-0.5% range (e.g., 0.05% in Germany or 0.4% in Italy). The degree of control in European countries is also high as in the UK. For example, in 2001 Switzerland knew with a high degree of precision that HIV-positive cases were distributed among homosexuals, IDUs and heterosexual with 25%, 15% and 50%, respectively. See Zurn et al. (2001) on the evolution of HIV/AIDS in Switzerland.

[6] The impact of AIDS in life expectancy of Low Income countries is visible relative to that in wealthy nations, both due to differentials in the number of HIV-positive cases and because wealthy nations have access to High Active Antiretroviral Therapy (HAAT). At the onset of the epidemic a HIV positive person was expected to develop AIDS and die within 6 years. Today, 6 years is the expected life of a HIV positive adult in a Low Income country. In wealthy nations, access to HAAT means that HIV-positive individuals can live normal lives and expect to be alive even 15 years after infection. Consequently, life expectancy in High Income countries is unaffected by the epidemic. An adverse outcome is that people in rich nations become complacent resulting in an increase in the number that become newly infected. (UNAIDS HIV/AIDS surveillance in Europe(2002), www.unaids.org)

[7] Prevention efforts in Uganda and Senegal have been significant since the mid 1980s with results showing epidemics kept under control. To talk about Sub-Saharan Africa does not imply that North African is free from HIV/AIDS. In Sub-Sahara estimates are due to data availability. North Africa (e.g., Egypt and Sudan) is believed to have a heterosexually driven epidemic with a least 2% HIV-prevalence, but lack of data and surveillance means that little can be said(UNAIDS, 2003).

the warnings and questioned the link between HIV and AIDS.[8] The consequence is that in 2004 South Africa's HIV-prevalence is estimated to be between 15% and 20%, i.e., a rate at least 150 times higher than in the UK, and, together with neighbouring countries (Namibia, Botswana, Zimbabwe, Zambia and Malawi), the highest in the world (UNAIDS, 2003).

Nowadays the only policy tool to keep the HIV/AIDS epidemic under control is prevention by means of wide spread information campaigns to effectively modify HIV risk-related behaviour. Furthermore, the campaigns have to be continuous to keep in alert those who are already knowledgeable, and to inform incoming younger generations before these engage on HIV risk-related behaviour. This is applicable to all countries in the world because in any country generation renewal and migratory movements implies an ever changing stock of knowledge in the population.[9]

Despite the biological complexity of the virus,[10] the social dynamics of HIV/AIDS are well understood. Poverty is a major underlying cause with wide ranging consequences. Poverty implies less

[8] Until very recently some African leaders believed that westerners who go on about AIDS in Africa are accusing black people of immorality. South Africa's president Thabo Mbeki complained that "…*we are blamed as criminals and seen as human beings of a lower order that cannot subject its passions to reason… as natural-born, promiscuous carriers of HIV, unique in the world, doomed to a mortal end because of our unconquered devotion to the sin or lust…*", see The Economist special survey on Sub-Saharan Africa, The Economist (January 2004), Vol.370,No.8358.

[9] Continuing with the UK example, the first campaigns of the 1980s were successful at controlling the epidemic, but the population became complacent over the late 1990s, while posterior less energetic information campaigns has resulted in the number testing positive for HIV slightly above the projected trend. What is significant is that such increase are among heterosexual that were too young in the mid 1980s to had been affected by the early campaigns (see UNAIDS Fact Sheet for Great Britain and Northern Ireland, 2002). Furthermore, since the late 1990s migratory movements have increased the number of HIV positive heterosexual in the UK due to cases acquired in Sub-Saharan African. For example, in 1999, 64% of heterosexual diagnosed had been infected in Sub-Sahara. This trends worries politicians in Western countries. The UK and Switzerland have proposed systematic testing of will be immigrants or asylum seekers(see The Guardian, 6 February 2003 & 13 February 2003, and the NZZ am Sonntag, 18 January 2004).

[10] One difficulty on finding a vaccine is that when virus are around for a while they learn to mutate. Worst for developing economies is that migratory movements implies a mix of HIV-subtypes that cause new drug-resistant

chances to attend school where preventive information might be readily available.[11] Poverty is also linked to high spatial mobility with temporary migrating workers, forced displacements and official government workers (e.g., soldiers) spending long periods outside their homes. For example, Janjaroen and Khamman (2002) find that an increase in workers from Myanmar and Cambodia is a significant factor in the recent upsurge of HIV-prevalence in the Thai population, while in 1996 studies in rural Kwa-Zulu Natal showed that migratory workers were 3 times more likely to be HIV-positive than those who remained at home (Cornia, 2002). This particular aspect of the epidemic has important implications for China, a country with an internal migration that involves a floating population of 100 million, and with a general population that has so far received limited information from its own health authorities.[12] Poverty is also associated with aspects such as unemployment, specially in densely populated cities where youths might be driven to the use of illicit drugs with low hygienic means that increase the chances of HIV infection.[13]

strains to develop, as has been observed in Congolese villages. This creates further divides between developing and developed regions, since in the latter the virus is more stable (Velazquez-Campoy et al.(2002)).

[11] Furthermore, De Walque (2001) has shows that education in itself acts as a factor that significantly reduces the probability of becoming HIV positive. One suggestion is that investing on education leads people to put greater value in the probability of staying alive in the future. The suggestion is that in regions where the HIV/AIDS epidemic is rampant, measuring returns to education in the form of income are an under-estimation because these ignore the positive externality that education has on reducing the probability of becoming HIV-positive.

[12] In China, a mix of poverty and government mismanagement of the epidemic has lead to a huge infection rate in some farming communities due to blood/plasma sales. A surviving strategy for struggling farmers was to sell their blood for $5 at a time. Blood of the same type was pooled and centrifuggated to extract the plasma, with the remaining blood cells injected back to the farmers. The result is that in the Henan Province only, 1 million are believed to be HIV positive, with some villages showing prevalence of 75%. In October 1998 the sell of blood was forbidden by the authorities, but this has not come about with the needed information to make people aware of the consequence of been HIV-positive, thus transmission continues been a worrying factor (Cornia, 2002).

[13] For example, in 2002 Indonesia had a population of 231 million, many suffering the consequence of Asia's economic slump of the mid 1990s. This has increased the number of young unemployed, specially in populated cities such as Jakarta that has experienced a 60% rise in the number of IDUs.

Clearly, poverty propagates HIV/AIDS widening the inequality gap between the rich and poor regions of the world, as well as worsening within-country income inequality.[14] However, there are factors other than poverty that also play an important role on the spread of HIV/AIDS, such as social costumes and norms, behaviour aimed at class-differentiation, or traditions (e.g., tattooing). For example, Thai men often visit sex workers in groups as a recreational activity after work. This traditional social activity in Thailand is known as 'Commercial Sex Patronage', and men from all social strata (from students to soldiers to white-collar workers) practice it. Estimated show that 95% of men visiting sex workers do so on a commercial sex patronage outing, with 70% of these declaring to participate as result of group pressure. What is worrying is that despite large scale HIV/AIDS information campaigns since the early 1990s, a 1995 survey of Thai men showed that as many as 55% of those participating in commercial sex patronage reported irregular use of condoms, and this in a country where HIV-prevalence among (male and female) sex workers are estimated to be as high as 30% (Vanlandingham et al (1995) and UNAIDS Fact Sheet on Thailand). In terms of behaviour that differentiates between social classes, it is well known that in many countries of Africa, Asia, South East Asia and the Pacific, the ability to pay for sex or to have many partner at any one time has often acted as a signal of power and wealth.[15]

---

[14] The economic implications of HIV/AIDS are clear since its major impact falls in prime age populations, thus, not only leaving orphans with lower parental and productive skills, but also instantaneously reducing the country's productive capacity. Some estimates point to a decrease of 1% GPD growth rate in low income countries with medium/high HIV-prevalence, while micro-estimates suggest that two-adults households loosing an adult to AIDS can experience a 30-40% drop in household income over the subsequent year, relative to those who maintain two adults (Cornia and Zagonari, 2002). Other figures question the macro-economic effect of AIDS. For example, in The Economist (November, 1003, Vol.369, No.8352) some argue that the 1% reduced growth does not account for the positive external effect that it has on leaving more resources for those that remain alive.

[15] De Walque (2001) shows that as Uganda's information campaigns take effect, individuals with more education and, on average, higher earnings, have systematically reduced the number of irregular partners. The result is a reversal in the gradient between education and incidence of HIV, i.e., pre-campaign estimated show that more schooling implied a greater (economic) ability to increase the number of partner, thus greater chances to become HIV-positive, whereas surveillance data in 2000 shows that those with more schooling have modified their behaviour to the extend that education is now negatively related to incidence of HIV/AIDS.

Whether the underlying causes for the spread of HIV are either poverty, socio-cultural norms or a combination of the two, modifying people's behaviour leads to a control of the epidemic. This is highly relevant for regions such as South East Asia and the Pacific, where the onset of the epidemic occurred with a time-lag of almost 10 years with respect to Europe, Africa and the American continent.[16] UNAIDS scientist place South East Asia and the Pacific as the second worst hit regions in the world, with HIV trends similar to those experienced in Sub-Sahara when this was at a similar stage in the evolution of the epidemic. However, the hope is that the number of HIV incidence will pick at a much lower HIV-prevalence than in Africa, given that various regional governments, joint with international organizations, have acted faster than in Africa, with expected positive future outcomes. For example, Cambodia is similar to Sierra Leone, with 20 years of internal conflict leaving the country in total devastation, no health infrastructure and a severe lack of skilled human resources. Since the first reported cases in the 1990s the epidemic grew to alarming rates reaching 4% in less than five years in urban adult populations, as experienced in war-driven areas of Africa. At reaching such HIV-prevalence, the Cambodian Ministry of Health has engaged on preventive measures such as a fierce condom use programme for the police force an soldiers. Although the campaigns are still weak and target only some sectors, the results from 1999 to 2001 is a drop of new HIV cases among government workers from 4% to 2.6% in the Phnom Penh area. Results like this encourage the believe that, relative to the African case, South East Asia and the Pacific will experience lower HIV-prevalence once the epidemic picks.

The fact that new emerging epidemics have occurred in South East Asia and the Pacific does not makes this region a special case.[17] The worry to scientist is that the region holds one quarter of the world's

[16] The first cases were reported in Thailand around 1985 in female sex workers. Until 1990 the region's HIV prevalence rate was well below that in the West. In Indonesia and Cambodia the first cases were in the mid 1990s. Since then, the HIV virus has spread at alarming rates, in some cases faster than in the Sub-Saharan Africa (e.g., Cambodia and Papua New Guinea). In Indonesia, only between 2000 and 2002 infection in high risk groups has gone from 0.1% to double figures, with 47% infection among IDUs in Jakarta, 26% among sex workers, and 22% among prisoners (WHO's HIV/AIDS Strategic Framework, SEA Region 2002-06, www.who.org).

[17] Emerging epidemics are starting to appear in Central Asia, the Baltic States and North Africa. (UNAIDS, 2003).

population, and, in most cases, very young populations vulnerable to infection due to the high correlation between youth and HIV risk-related behaviour.[18]

A further characteristic of this region is a wide social, economic and demographic diversity. Average HIV-prevalence in the region is 0.8%, but country-specific prevalence range from 0 % (Mongolia) or 0.1% (Australia or Hong Kong), to as high as 2.8% (Cambodia, followed closely by Myanmar and Thailand). Regional countries also vary widely with respect to wealth (e.g., 2001 per capita GDP range from $320 (PDR Laos) to $34,500 (Hong Kong)) and development levels (e.g., the 2001 UN Human Development Index ranks Australia as the 4th most developed country in the world, while its neighbour Papua New Guinea is ranked 132 in a scale of 162 ranks). Such diverse range of wealth and development is highly correlated to the causal focus of the spread in each country. Thus, whereas poor countries have seen the emerging epidemic been driven by commercial sex and among intravenous drug users (e.g., Viet Nam, Indonesia or Cambodia), wealthier countries (e.g., Australia and Singapore) or territories such as Hong Kong, can pin-down heterosexual contact between young people or MTMT as the main focus of HIV-infection. There are also exceptions such as Mongolia's zero HIV-prevalence, mostly due to the fact that it has had little influence from the outside world. Although it opened up as a democracy in the mid 1990s, the nomadic style of its people implies almost no transfers of the virus, with 3 reported cases since 2001, two acquired by sexual contact with foreigners.[19] The same is not true for China, where the 0.1% estimate reflect a huge population size with a hidden emerging epidemic. The Chinese problem has its focus on the share of contaminated needles by IDUs and unsafe blood-plasma collection, while heterosexual transmission becomes a follow up reason to the other two. On the other hand, a major transmission factor in the five neighbouring countries of Myanmar, PDR Laos, Thailand, Viet Nam and Cambodia is related to sexual practice behaviour (the use of sex workers) and a very high level of population mobility, thus heterosexual transmission is the main problem, specially

[18] In populations of countries with the highest HIV-prevalence in the region, namely Cambodia, Thailand and Myanmar and Papua New Guinea, the median age is 19, 29, 21 and 19, respectively, figures that contrast with countries such as Australia where the median age is 34. In other regions of the world with emerging epidemics the median age is significantly above that of countries in South East Asia and the Pacific Region(e.g., 35 in Russia).

[19] Source: UNAIDS National Response Fact Sheet: http://www.unaids.org/nationalresponse/print_html.asp.

among female sex workers and their clients. For Thailand, intravenous drug users is also an important focus that rivals the sex industry; with HIV-prevalence in the adult population at around 2%, this goes up to 50% among IDUs in Bangkok, a rate that approximates that of female sex workers in the same city.[20]

In 2003, most countries in South East Asia and The Pacific had national strategic frameworks to inform the general population on the social dynamics of HIV/AIDS. These campaigns are often government initiatives backed up by international organizations (e.g., UNAIDS or UNICEF).[21] It would be safe to believe that such campaigns have promoted information in the region at least since the early 1990s (e.g., Thailand and Papua New Guinea) and at the latest in 1998/1999 (e.g., Myanmar and Mongolia). Such campaigns can be thought as effective policy tools if they spread information to all sectors in the population, so that individuals are aware of the transmission mechanisms behind HIV/AIDS, on the benefits of regular testing and on any other practice that reduces exposure to the virus. For example, if information reaches rural areas, spouses of returning migratory workers can themselves avoid infection by choosing low risk behaviour if they understand the consequence of their partner's possible HIV high-risk behaviour.

Thus, campaigns that reach all sectors in the population could act as effective instrument to reduce (within-country) health as well as wealth inequality in the long run. The opposite is also true: if information campaigns act only on those in a position to benefit from them, social inequality will widen. This might be the case, for example, if such campaigns were to focus on children who attend school, neglecting those who are not able to do so.[22] Some evidence in Africa suggest that households with adults that carry the virus halved

[20] Source: AVERT, www.avert.org/aidssoutheastasia.htm

[21] See the Regional and Country specific Fact Sheets in www.unaids.org for a full systematic account of the main strategies within each country, total expenditure and status of the campaign. This Fact Sheets with information are updated regularly.

[22] The region of South East Asia and the Pacific has countries that account for the lowest levels of school enrolment in the world. For example, in Papua New Guinea, Myanmar and Cambodia a quarter of primary school age children do not attend school, while secondary school attendance of the corresponding age group does not reach 30% in places such as PDR. Laos, and only 50% in the Philippines. The opposite side of this coin is that many children become full-time workers at a very early age (e.g., 45% of child labour in East Timor, and 22% in Myanmar or PDR Laos (ILO Yearly Statistics Book, 2002)).

the spending on their children's education as compared to households of similar status but with no HIV carrier (The Economist (January, 2004), vol.370, No.8358, on a study of Côte d'Ivoire), effects that are often worst for girls in the household. Thus, information campaigns need to target all social and economic groups to stop the self-propagation of the HIV/AIDS epidemic among those who are more susceptible to HIV infection because of their social condition. Furthermore, targeting children's awareness is important because these are individuals in the population who are – in its vast majority – not yet sexually active and/or not yet using intravenous drugs, meaning that information campaigns can help modify their behaviour before they start to face risk-choice situations with respect to HIV infection. This is specially true in a region where, as scientist suggest, the greatest factor to fear is the youth of its population alongside traditionally rooted high-risk behaviour with respect to HIV infection.

The following section describes the dataset analysed in this paper. A positive factor in this dataset is that surveyed countries are representative of Western-style (High Income) democracies, Communist-style administrations, and countries with freely elected governments but best classified as Low Income economies. This division allows the study of differentials on HIV/AIDS related knowledge of children who live in different degrees of democracy and free-press, and, therefore, an evaluation of Amartya Sen's proposition on the effect of a country's freedom on the evolution of HIV/AIDS (see Sen (2002), Peterson (2000), Cornia (2002) and Manning (2002)).

3 A UNICEF Children Survey in East Asia & The Pacific

During the first half of 2001, UNICEF regional offices in South East Asia and The Pacific region, in co-operation with UNAIDS, conducted an opinion survey on a sample of children from 14 countries and 3 territories of South East Asia and the Pacific. The aim of this survey was to obtain an objective and independent understanding on children's views, opinions and actual knowledge with respect to a wide variety of issues. Using door-to-door interviews, a total of 10,073 children with ages between 9 and 17 were interviewed by an experienced team, selected and co-ordinated by Research International (Thailand), a Bangkok based research consulting institute. It would be difficult to obtain a truly representative sample from each country considered in the survey:[23] the final sample is a selection of children who are within-country representative with respect to age-group, gender, urban/rural composition and socio-economic level (as defined by income groups for a given city or province). The selection system is not dissimilar to that of many well known household surveys; a household containing children on the required age group is approached and the household head (either a parent or adult guardian) is asked for permission to interview the child. The adult can refuse or accept. In case of refusal the interviewer will keep on asking households with similar (sample representative) characteristics until a household within the required sample-strata agrees to have the child being interviewed. In this case, the adult is asked to respond to a set of questions at household level (household size, occupation of main income provider and total household income). Following this, the interviewer asks to be left alone with the child for approximately half an hour and until all questions in the survey are responded by the child. With this system all answers are assumed to be free from adult's influence. If the household has more than one child in the required age-range only one child is selected for the interview, with the random selection based on a the Kish Grid method.[24]

In total, there are some 40 questions in the survey, including questions at household level. Questions answered by the children could be classified into categories such as schooling status (goes or does not go to

[23] The 14 countries and three territories listed in this section vary widely with respect to most economic, social and demographic indicators. The region itself contains one quarter of the world's population. A representative survey of children in such densely populated area would require a mammoth task outside the scope of any survey. We view the 10,073 sampled children as providing representative opinions with respect to regional children.

[24] Example of similar selection strategies are found in household surveys such as the German Socio Economic Survey (GSOEP), or the Spanish Enquestas de Poblacion Activa (EPA); they also select households with respect to locations that define a particular socio-economic group. If on approaching such household there is no response or refusal to interview, the interviewer will select a household with the same location-characteristics. The assumption is that the replacement household is equally representative within the sample. In the case of the UNICEF survey one could assume that parents/guardians that accept to have their children interview might lead to a selective sample with respect to the characteristics of the child's views and understanding. However, it would be difficult to estimate and unbiased distribution of children's views and understanding with respect to the issues explored within the survey. Thus, we assume that refusal to respond is not related to such issues, and at the same time, we can think of the results as been representative of the better off within each social, demographic and economic class. See Appendix A for a exploration of issues that might affect the representative quality of the sample.

school) and opinions referred to school for those who do attend, questions on substance abuse (e.g., smoking, alcohol drinking, sniffing glue), understanding of their own country's socio-politic situation, relationship with adults within their surroundings, level of happiness, satisfaction and relative future expectations (e.g., if they expect to be better off than their parents or hopes for future occupation), knowledge on various aspect on the Human Rights of children, and questions that measure the wealth of the household by means of possessions (e.g., the existence of electricity or tap water in the household). A final section aims exclusively at eliciting objective information on the child's knowledge with respect to issues concerting HIV/AIDS. All surveyed children answer to all the sections, except for the final section (on HIV/AIDS knowledge) that is answered only by children of age 14 or older.

The aim in this paper is to find out if geopolitical, social or economic factors are significant at determining differentials in children's capacity to avoid HIV-infection. This is motivated by the assumption that information campaigns can only be effective policy tools if they target and reach the population as a whole and, in terms of children, if they are able to equip them with the necessary knowledge to avoid infection and/or modify their health related habits before they become young adults. Therefore, the main block of information needed for analysis is provided by the 4,630 children who are 14 years of age or above, since only these answer to the final section in the survey.

Each of the 4,630 children belongs to one of 14 countries or 3 territories (East Timor, Hong Kong and Macao) in South East Asia and The Pacific Region. Table 1 shows the distribution of the sample between the 17 States, alongside a selection of key comparative macro-variables. Approximately 40 % of surveyed children in each country are between 14 and 17 years of age. All countries and territories are classified in one of three categories according to income and (recent past or present) government style. The Western Style/High Income countries include two territories (Hong Kong and Macao), while the Low Income countries include the other (East Timor). Countries such as Malaysia and Thailand are border cases with respect to income: Malaysia has a low per capita income relatively to others in the Western-style group, but a relatively high Human Development Index justifies its categorization. On the other hand, Thailand has an income two times higher than the next Low Income country (China) and is only relatively far from Malaysia's GDP, but its HDI is more than 20 rankings below Malaysia, thus, it becomes part of the Low Income group. Communist style economies might not be straight forwards under classic communist regimes. For example, Viet Nam and China are slowly opening up to capitalism, while Mongolia's political organization is close to tribal-nomadic societies, not too dissimilar to the social structure of Papua New Guinea. Nevertheless, years of communist style regimes, either at present or in the recent past, implies that such countries might experience similar processes in the diffusion of information, for example, information relating to sexually transmitted diseases as is the case of HIV/AIDS. Amartya Sen, David Patterson and other prominent development theorist[25] provide theoretical arguments suggesting that authorities in a democratic system will have more commitment on addressing the epidemic, with poverty per se not being the only factor leading to higher HIV-prevalence. For example, in a democracy those in charge feel more accountable to help the population in dealing with epidemics because inaction is often punished by being voted out of office. Likewise, democracies may have better channels to diffuse information. The division of countries between High Income democracies, Low Income democracies and non-democratic (Communist style) countries allows for a evaluation on Sen's (2002) proposition.

Table 1 shows the huge diversity between countries in the region with respect to development, wealth and of course, HIV-prevalence. Even within each group there are remarkable differences: Hong Kong has a per capita GDP eight times greater than Malaysia, the development indicator HDI shows Australia as one of the most developed countries in the world while East Timor, Papua New Guinea and Cambodia are just a few ranks away from Sierra Leon (the lowest ranking in the world). Finally, HIV-prevalence varies between countries in a way that is difficult to pint point either as consequence of wealth, regime style or development.

[25] See Manning (2002), Peterson (2000) and Sen (2002).

Table 1: Surveyed Countries, Distribution of Sample (n=4,630) between Participating Countries, and Selected Macroeconomic Indicators

COUNTRY	Sample Size in [14,17] age category	HIV/Aids Prevalence in 15-49 age range	Median Age of the Population	Per Capita Gross Domestic Product	Rank in Human Development Index [HDI]
Western Style Economies					
AUSTRALIA	**222 (44.4)**	0.10 %	34	$ 19,100	4
HONG KONG	**134 (44.7)**	0.10 %	35	$ 24,300	26
MACAO	**83 (41.4)**	0.05 %	24	$ 14,100	25
MALAYSIA	**261 (50.7)**	0.40 %	21	$ 3,700	58
SINGAPORE	**237 (47.2)**	0.20 %	31	$ 20,700	28
SOUTH KOREA	**278 (55.6)**	0.05 %	30	$ 15,100	30
Low Income, Non-Communist Economies					
EAST TIMOR	**188 (36.7)**	0.05 %	18	$ 520	143
INDONESIA	**467 (45.2)**	1.00 %	20	$ 680	112
PHILIPPINES	**198 (39.6)**	0.10 %	22	$ 910	85
P.N.GUINEA	**251 (50.2)**	0.70 %	19	$ 550	132
THAILAND	**228 (45.6)**	1.80 %	29	$ 1,880	74
Communist Style Economies					
CAMBODIA	**213 (41.9)**	2.80 %	19	$ 280	130
CHINA	**860 (45.3)**	0.1 %	29	$ 920	104
MONGOLIA	**265 (53.0)**	0.00 %	22	$ 420	117
MYANMAR	**262 (52.4)**	2.20 %	21	$ 420	131
PDR. LAOS	**215 (43.0)**	0.05 %	18	$ 320	135
VIET NAM	**268 (44.7)**	0.30 %	20	$ 420	109

Note 1: The percentage shown next to each country's sub-sample is the percentage relative to the total number of children surveyed in that country. For example, out of the 10,073 children, 497 are from Malaysia. Of these, only 261 are in the age group [14,17]. Therefore, the 261 children represent 50.7% of the total number of children surveyed in Malaysia.

Note 2: The region's HIV-prevalence is estimated at 0.8 %. As a comparative measures, Europe's HIV-Prevalence is estimated at 0.3% with countries such as the UK recording low prevalence (0.1%) relative to high prevalence countries such as Spain & Switzerland, each with equal rates at 0.5%. All HIV prevalence are based on 2001/2002 estimates from UNAIDS Country Fact Sheet (www.unaids.org).

Note 3: Gross Domestic Product is in USA Dollars at current 2001 levels. Source: The World Bank (www.worldbank.org).

Note 4: The Human Development Index (HDI) is a measure of the country's development achievements that takes into account life expectancy, educational attainment and adjusted real income. It is developed by the United Nations Development Program (UNDP), and it ranks all UN Member States in descending order according to their relative position. The ranking presented in Table 1 refers to 2001, where Norway was ranked 1 (highest development level) and Sierra Leon was ranked 162 (lowest development level). (www.undp.org)

Note 5: The median age is taken from ILO projections, based on latest population census. See Appendix A for more information on macro-measures.

The UNICEF survey is indicative of information on a cross-section of households for each of the countries in Table 1. Section 3.1 explains and motivates the construction of the main outcome variable of interest, namely, an individually assigned probability of being badly equipped to prevent HIV-infection. This is constructed with each child's answers to questions on HIV/AIDS knowledge, indicators of risky health habits and country's specific HIV-prevalence. Section 3.2 describes the geopolitical, social and economic indicators identified in the survey, and provides details on household and children's background variables employed in the estimation process.

3.1 Constructing an indicator for the probability that a child is equipped to avoid HIV-infection

One aim in this paper is to construct an indicator that best approximates the chances a child has to avoid becoming HIV-positive when faced with a high risk situation with respect to HIV infection.[26] Section 2 demonstrates that knowledge on HIV-transmission and/or safe behaviour are the best available tools for people to avoid infection. The survey provides two questions highly indicative on each child's degree of knowledge on HIV/AIDS. In the first question the child is read the following statement:

> *Question Q31:*
> *'We would like to get some idea of what you know about how HIV/AIDS is transmitted from one person to another. In your opinion, how is HIV/AIDS spread between people?*

The interviewer is specifically indicated to allow the child to answer as many possibilities as the child perceives to be the correct answer, so that for each child the answer is a record of at least one, but in most cases, multiple answers. These can range from 'kissing someone that is HIV-positive', to 'been bitten by a mosquito that has bitten someone who is HIV-positive', to 'Through unprotected sex' or 'because does not use condoms when having sex'. Known to date, the possible causes leading to a HIV-positive status are all related to human contact with HIV-infected human body fluids. With this, 4 groupings of potential causes are best understood and targeted by information campaigns, namely (1) *through unprotected sex*, (2) *sharing needles among intravenous drug users*, (3) *from mother to infant either before or during birth*, and (4) *through any other blood-transfusion related situation* (other than sharing needles). The 4,630 children in the survey can be classified according to their awareness on how many of these four potential situations cause HIV/AIDS, with 38% who do not know any,[27] while 25%, 25%, 10% and 2% answer by identifying one,

---

[26] The HIV status of these children is unknown. The analysis starts by assuming they are HIV-negative. With this, the implication is that in the event of facing a situation of HIV high-risk, their knowledge and risk behaviour would be the best factors to predict the risk of infection.

[27] The 38% are either those who go directly for a 'don't know' answer, or those whose multiple answers are completely incorrect, such as 'touching a HIV-positive person', 'sharing food with someone that is HIV-positive', etc. Those who answer 'by a mosquito that has bitten someone who is HIV' are taken as correctly identifying that

two, three or all four correctly, respectively. The survey provides a second question that gives evidence on the knowledge of children on HIV/AIDS:

Question Q32:
'How do you know if a person is HIV positive?

Again, the interviewer is specifically indicated to leave the child to answer freely to as many multiple possibilities he or she might feel provides a correct answer. Examples of answers given by children to Q32 included 'he/she has dark skin', 'he/she looks sick', 'he/she has spots', or 'by a blood test', this latter been the only possible correct answer. Among all children, only 22% identified 'a blood-test' as the correct method to know if a person is HIV-positive. Using these two questions (Q31 and Q32 in the survey), it is possible to construct a categorical variable (k) that assigns a 5 to children whose answers include all four causes of transmission and, at the same time, are able to identify a blood test as the only valid measure to know a person's HIV status. Those who score 5 (i.e., $i \in n, s.t.\ k_i = 5$) would indicate superb knowledge relative to those with no idea as to how HIV is transmitted and, furthermore, fail to identify a blood test as only true indicator of HIV status: a child with knowledge like this can be thought as scoring 0 in our knowledge scale (i.e., $i \in n, s.t.\ k_i = 0$).

Besides HIV/AIDS Knowledge, the survey is also indicative on the probability that a child engages on high-risk health-related behaviour.[28] The question is posed as follows:

'blood transfusion' is a potential transmission cause. Although those who answer with 'something', no matter how incorrect might show – at least – awareness of the virus as compared to those answering 'don't know' straight away, we still classify all in one group to reflect that no knowledge or incorrect knowledge is equally 'un-preventive' in the case of facing a high HIV risk situation. The assumption is that they have to know how to keep safe from the virus rather than just knowledge that the virus exists.

[28] As the question shows, the children are not asked directly about their own drug-taking status. This strategy can help reduce the well known condition that children will fear reprisal on declaring their own habits and/or preference to be seen positively by the interviewer. However, if a child declares to know well someone of his or her own age that smokes or sniffs glue, there are high chances that the child has engaged on similar acts. However, this not being the case, what is true is that young children move in environments that define their own socio-economic status. Thus, the answers given by this children are proxies for either children that engage on high risk taking

Question Q5a/Q5b:

'Have any of your friend or anyone you know well, of about the same age as you, ever tried the following substances even if only once, or have become addicted to them?(Each substance is read independently and the child answer YES, NO, DON'T KNOW)

'Smoking'

'Alcohol'

'Sniffing glues or other chemical'

'Illegal drugs'

'Abuse of prescription drugs or other medicines'

Using Question Q5a/Q5b, the survey allows for the construction of another categorical variable (r) that classifies children according to their exposure to high health-risk situations. If a child $i \in n$ answers indicating contact with children who take part in all 5 health hazards we allow $r_i = 5$, while a child $i \in n$ indicating no relation to any of the 5 health hazards scores such that $r_i = 0$. The sample is distributed so that 33% score $r = 0$, 24% score $r = 1$, 31% score $r = 2$, 8% score $r = 3$, 3% score $r = 4$ and 1.4% score so that they know people involved in all 5 health risk situations ($r = 5$).[29] Appendix B provides summary statistics for both k and r, allowing for different sub-populations as defined in Table 1.

The paper evaluates if geopolitical and socio-economic effects explain probability differentials on becoming HIV-positive. This probability depends on local factors faced by each individual (e.g., HIV-prevalence in the region), but it is well demonstrated in Section 2 that more importantly it depends on the individual's capacity to avoid infection, where 'capacity' is the direct result of each individual's risk attitude and knowledge, with both factors closely interlinked. Clearly, it would be difficult to ask children directly 'what is the probability that you will avoid HIV/AIDS once you enter adulthood?'. However, using the two constructed categories k and r, it is possible to define a variable indicating the following: If children, on entering adulthood, were all to face an identical situation with respect to risk of HIV infection, what would be

behaviour who have the child's knowledge or household attributes, or, the answers declare a high probability that the same child who answers is declaring his or her own health habits.

[29] The survey is further informative with respect to how knowledgeable are children with respect to condoms and with reference to using condoms. This questions are kept aside as indicatives to be studied separately in future version of the paper.

the probability that they are badly equipped (with respect to knowledge and health risk behaviour) to avoid infection?", in other worlds "What is the probability that the state of HIV/AIDS knowledge joint with the chance of been engaged in high-risk health behaviour leads a child to become HIV-positive in the future?". Let this probability be named P_A. Assume the a sample contains two children ($n = 1,2$) who answer with combination $(k_1 = 5, r_1 = 0)$ and $(k_2 = 0, r_2 = 5)$, respectively. I have two assign a value of P_A to each child where the two values indicate their (within-group) relative position. Clearly, $P_{A,1} \leq P_{A,2}$, for example, $P_{A,1} = 0, \quad P_{A,2} = 1$ since the (underlying and unknown) probability that child number 1 becomes HIV-positive, relative to child number 2, is much lower. For a given set of combinations (k, r), we end up with $P_A = P_{A,1},.....,P_{AN}$, all in the [0,1] interval, where each child is given a relative probability value, with low values indicating low chances of becoming HIV-positive given knowledge and risk behaviour, and high values, the opposite. It is a ranking over an arbitrary interval [0,1] that I choose to interpret as a 'within sample potential relative probability of becoming HIV-positive'. Two points need to be made. First, these children are very close to becoming young adults, with 4% of them already indicating sexual activity.[30] Therefore, we assume that acquired knowledge at present is very indicative of future knowledge that might not change significantly, specially for those in the older age group (16 or 17), and, at the same time, heath risk behaviour can be thought as changing at a slower pace than knowledge.[31] The second important factor is the timing of the information campaigns. We are analysing children born between 1983 and 1987. These children have been young teenagers/adolescents for at least two years, so we are dealing with children who are aware of their social surrounding. At the same time, information campaigns have been going on around in the region since as early as 1989 (Thailand and Papua New Guinea), to the latest initiatives in Cambodia (late 1990s). Thus, it is correctly assumed that P_A reflects the effectiveness of these information campaigns. Through the use of appropriate econometric tools, finding differentials on P_A by either geopolitical or socio-economic treatments

[30] The survey has a section that ask children about knowledge an use of condoms. These questions will be used more extensively in future versions of the paper to test for differentials in knowledge between groups of different risk.

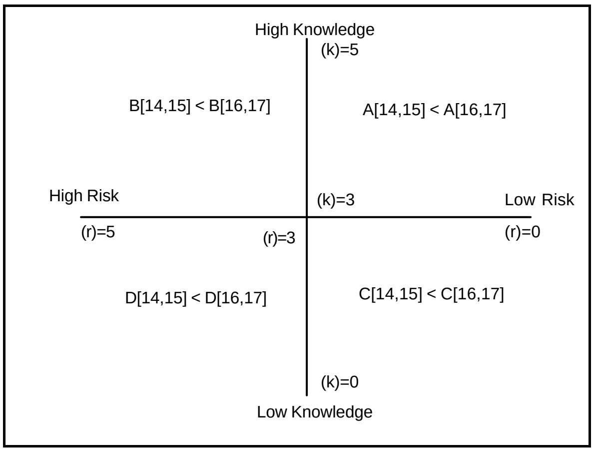
[31] The idea is that addiction is more difficult to give up that for knowledge to come along (need reference on addition here).

might indicate the working effectiveness of such information campaigns (see Section 4 for an account on the econometric methods). In the absence of an index P_A that unifies both k and r an applied econometrician might be inclined to examine each of the categorical variables with some nonlinear categorical tool, for example, an ordered response model, or in the case of joining k and r into all possible combinations, one could also use an interval regression type of model. These two approaches are often restrictive in their functional forms, with underlying assumptions (e.g., symmetry and normality of the error terms) often difficult to justify, while estimated coefficients in ordered response models are often questionable in their interpretation. Building an index P_A allows for a clear interpretation of the variables k and r, with no other assumptions that those implied by 20 years of empirical evidence on the evolution of the HIV/AIDS epidemic.

Classification of the 4,630 children in the [0,1] interval can be done in multiple ways. A simple one would be to classify them in ascending order by (k,r) combinations, allowing values in the [0,1] interval to interpret their relative positions. This seems artificial, as it breaks the [0,1] interval into 36 discrete sub-intervals. Furthermore, as has been suggested, knowledge and health risk are important determinants for the probability of HIV infection, but so is the child's country specific HIV-prevalence, since this latter proxies for 'random shocks' to the probability of infection. For example, a child becoming a young adult might have an accident leading to the need of a blood transfusion; if so, it is the country's specific infection rate, as opposed to knowledge or behaviour, that will dominate the probability. Thus, we need to allow for each i^{th} child in n to be given a value $P_{A,i}$ that reflects this 'country specific random shock'. To this aim, this paper adopts the Beta distribution as an intermediary devise to draw values on the [0,1] interval that reflect $(k_i, r_i, HIV_{i \in country})$ for each i^{th} child in the sample of size n. The beta family $B(a,b)$ allows for distributions of random outcomes in the [0,1] interval, where the parameters a and $b-a$ define location and scale. For $a=b$, the distribution is symmetric with mean and variance determined fully by a,b. Otherwise the distribution is asymmetric. To construct P_A using the Beta distribution, first we classify each child in one of four quadrants according to their combination (k,r). Table 2 illustrates this. In terms of avoiding infection, a child in the quadrant A is better equipped than a child in any other quadrant. Assuming that knowledge is a more important factor than risk,[32] a person adverse to becoming HIV-positive would rank her quadrant position as follows: $A \succ B \succ C \succ D$.

Table 2: Summary, dispersion of sample by Knowledge (k) & Risk Behaviour (r): Determinants of P_A .

Quadrant Distribution of P_A



Distribution of sample n=4,630 by Quadrant and Sub-sample definitions.(percentages with s.e in brackets)

	Full Sample	West. Style	Comm. Style	Low Inc. C.
% Q.A				
[14,15]	6.2 (0.4)	6.2 (0.7)	5.7 (0.6)	10.4 (0.7)
[16,17]	10.0 (0.4)	8.6 (0.8)	4.5 (0.7)	10.5 (0.7)
% Q.A				
[14,15]	1.1 (0.2)	2.0 (0.4)	2.3 (0.4)	1.2 (0.2)
[16,17]	2.3 (0.2)	2.1 (0.4)	1.7 (0.3)	1.5 (0.3)
% Q.A				
[14,15]	38.9 (0.7)	37.3 (1.4)	36.6 (1.3)	37.4 (1.1)
[16,17]	32.7 (0.7)	32.5 (1.3)	34.4 (1.3)	35.5 (1.0)
% Q.A				
[14,15]	3.5 (0.3)	4.9 (0.6)	7.2 (0.7)	1.9 (0.3)
[16,17]	5.4 (0.3)	6.5 (0.7)	7.7 (0.7)	1.7 (0.3)

In the construction of P_A , each child's value is drawn randomly from a Beta distribution that reflects his of her quadrant position in Table 2. For example, a random draw a^1 from $B(1,19)$ has an expected value of 0.05 with a variance of 0.0023, whereas a draw b^1 from $B(5,12)$ has expectation 0.30 and variance 0.01. Thus, when assigning a P_A value to a children in quadrants A, relative to another in quadrant B, the child in quadrant A would be given a P_A value $a^1 \in [0,1]$ drawing from $B(1,19)$, whereas the other child would draw $b^1 \in [0,1]$ from $B(5,12)$. The values of P_A are decided by such method, but (i) further distinguishing two sub-quadrants in each quadrant A, B, C, and D,[33] and (ii) for each of these sub-quadrants, older children

[32] Having the correct knowledge under a risky situation is best than knowing nothing at all because it implies better choice to avoid infection.

[33] The ideas is to imagine that each of the quadrants A, B, C and D are the 2-dimension plane. Groups of children with a combination (k, r) that places them in the same quadrant can be further divided into sub-quadrants that also reflects their relative position. This allows for a more refined partition of the outcome variable.

(i.e., 16 and 17) draw values from Beta distributions with a higher mean than younger children with equal (k,r) combinations: the interpretation is that older children will enter adulthood sooner, so that for an equal level of knowledge and risk behaviour, younger children are less likely to become HIV-positive, i.e., younger children have a bit more time to catch up knowledge or change behaviour.[34] Appendix C, Table C2 shows the 16 choices of Beta distributions, the corresponding (k,r) combinations for each choice and the theoretical mean and variance for each Beta specification. Furthermore, a country's HIV-prevalence can also be interpreted as a draw from a Beta distribution if one allows for the appropriate parameters (a,b). For example, Cambodia has a HIV-prevalence of 2.8% , where 0.028 is the expected value from a Beta $B(8,285)$. Specifying a Beta distribution that best approximates a child's country HIV-prevalence, a vector of outcomes P_A is constructed so that for each $i \in n$,

$$P_{A,i} = B(a,b|r,k) + B(a,b|HIV_{i's\,country}) \qquad \text{(PA.1)}$$

Expression (PA.1)is the explicit form of a more general function that defines $P_A = f(k,r;e)$. The assumption is that e, the country/environmental-specific random shock, enters additively. For each child's fixed combination (k,r) and his or her HIV-country prevalence we generate 200 random draws, thus generating a distribution of possible values P_A . Each of the 200 vectors has 4,630 elements and is treated as an outcome among 200 possibilities. Figures 1 and 2 plot the resulting 200 densities of the outcomes, and the average outcome of the 200, respectively.[35] The summary statistics next to Table 2 showed a cumulative probability of belonging to quadrant A equal to 0.162. The average of the simulations leads to a simulated probability of 0.135, thus reducing the chance of being in the best quadrant. Such reduction can be due to the

[34] Table C2a in Appendix C shows that at top levels of knowledge and lowest risk behaviour, values of P_A for children of any age draw from the same Beta distribution.

[35] The value 200 is based on a sensitivity analysis allowing for a random draw of 100 children. Trying out 25, 50,…,100, 110, and 120. After 100 draws the mean of the distribution is not significantly different. The choice of 200 seems as a save choice for simulating a distribution of probabilities for each sampled unit. The choice of 1000 units in the sensitivity analysis is more time efficient. Figures 1 and 2 define P_A with the mnemonic PHIVAL.

added ‘positive shock’, i.e., the country-specific HIV-prevalence (positive in the sense that it adds to the simulated probability of becoming HIV-positive). With P(Quadrant A)=0.135 and a median of the assigned probability placing 50% of the children in quadrants C or D, the overall picture is that children in the region are not well equipped to deal with the HIV/AIDS epidemic. This conclusion is based on the thought that a well prepared set of soon to be adults would be distributed such that at least 50% are in quadrant A or its proximities. Appendix C compares densities of P_A for different country blocks, and for each of the 17 countries (or territories). Comparing these provides a first view on geopolitical differences on P_A. Communist countries have a greater mass of children with simulated outcomes in quadrant A, although the estimated medians for Western, Low Income and Communist countries do not differ between them, meaning that for any block definition more than 50% of the children are well into the region of ‘poorly equipped to deal with HIV-infection’. Country comparison suggests that Viet Nam, Cambodia and Singapore are the three countries with greater density in quadrant A, each with medians well below the 0.60 overall sample median. For example, in Viet Nam 40% of children are in quadrant A. On the other hand, The Philippines, East Timor and South Korea show very low densities in quadrant A: in the Philippines, the probability of finding a child in Quadrant A is almost zero, while, in each case, the medians are well to the right of the average for the complete region, thus significantly increasing the number of children in quadrant D. Clearly, children in the Philippines are the worst prepared to deal with the epidemic resulting from a combination of low levels of HIV/AIDS knowledge and high health risk behaviour. What is more striking is that the positive random shock acts as a right-shifter for countries with the highest HIV-prevalence such as Cambodia and Thailand, yet, both countries have densities in quadrant A and P_A medians that places their population of children at much better position to deal with the epidemic, much better even than advanced Western style economies such as Australia. These might be first indications on the effect of advanced epidemics (as opposed to newly emerging ones) on both the general knowledge of the population and the knowledge of future generations in the population.

Figure 1: Density Estimates, 200 Draws

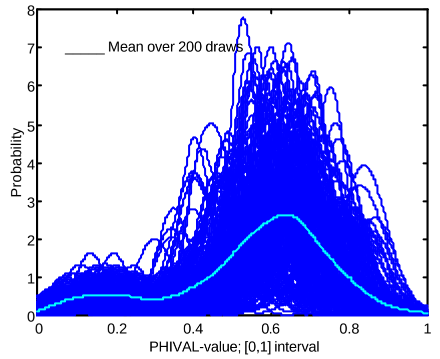
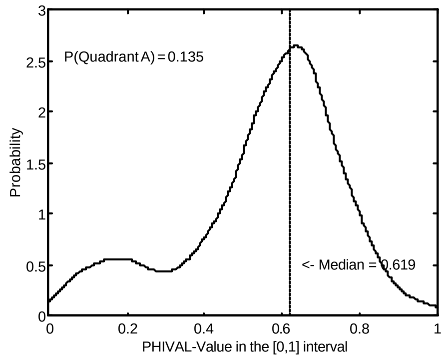


Figure 2: Density Estimates, Mean values



The above analysis is based on estimates that do not control for specific socio-economic variables that might affect how we compare differentials of P_A between and within countries. Section 3.2 presents a set of treatments to be considered in the empirical analysis. Section 3.2 also presents conditional variables that help sub-divide the population among comparable socio-economic 'cells'. This is the basis for this paper's estimation approach as explained in Section 4.

3.2 Description of Countries, Selected Effects and Background Variables

This section describes a set of treatments that are thought to have an impact on P_A, each identified by indicators of a geopolitical or socio-economic nature. Geopolitical effects means that the treated belong to a country or a group of countries. Socio-economic effects are defined by conditions in any one country. With respect to the latter set, the information in the survey limits the selection to urban households (versus rural), going-to-school status (versus no attendance)[36], been a boy, three relative levels of Socio-Economic Status

[36] This treatment is specifically defined as 'Going-to-school' because the survey provides no other information about the child's schooling status. The child is asked 'Do you go to school?' and answers 'yes' or 'no'. All children that go to school are at secondary level. I specifically define the control group as 'No school attendance' because, declaring not to go to school identify various reasons (e.g., no schools around the area, no money at home, has to work instead, etc). It would be interesting to estimate the effect of going to school against each possible alternative, but this is not possible because of a very small sample size for the controls in each alternative.

(SES),[37] and three (parental) occupational skill level (unskilled, medium skill, and skill/profession). Table 3 presents all treatments considered, with the corresponding partition between treated and untreated (control) samples, and in each case, the sample size.

Table 3shows seven geopolitical treatments. For example, the effect of Western style economies on P_A relative to 'All other countries', implies that the treated live in a Western style economy and the controls in a Low Income or Communist style economies, whereas the effect of Western style economies on P_A relative to 'Communist economies' implies that controls are children only from Communist-type of countries . The treatment 'High HIV-prevalence' versus 'Low HIV-prevalence' takes the cut off point at 0.4% because Table 1 seems to suggest this as a natural differentiating point between countries.[38]

[37] Household income is given by range cards that differ between countries, making it difficult to compare. Income can be specific to countries and, within these, to cities, by purchasing power. The SES level is built by Research International Asia (Bangkok), providing a relative index of economic well-being within and between countries, taking into account income and city of residence. Income nonresponse is low (3.8%) and only occurs in Australia, Cambodia, China, South Korea, Mongolia and The Philippines. The missing values are imputed by Mahalanobis distance (Rubin, 1987). The variables country, city, urban, work activity, household size, number of kids, school status and wealth items are used as background variables to impute the missing income values.

[38] Anticipating to results in Section 5, for any given treatment, the estimates of the propensity scores show densities that overlap over the [0,1] interval. This implies very good common support conditions. Thus, it is assumed that estimates of the effect of '*A*' on P_A with a given control group '*B*', assumes that the complementary treatment (the effect of '*B*' on P_A with a given control '*A*') gives the opposite results. Thus, we do not estimate alternative (*A*,*B*) combination to those proposed in Table 2.

Table 3 Presentation of Geopolitical and Socio-Economic Treatments, by Treatment and Control Group (number in bracket is sample size)

Between Countries		Within Countries		
Geopolitical Treatments (sample nt)	**Controls (sample nc)**			
Western Style Countries (nt=1,215)	All Other Economies (nc=3,415)			
	Communist Style (nc=2,083)			
	Low Income Countries (nc=1,332)			
Communist Style Countries (nt=2,083)	All Other Countries (nc=2,547)			
	Low Income Countries (nc=1,332)			
Low Income Countries (nt=1,332)	All Other Countries (3,298)			
High HIV-Prevalence $\geq$ 0.4% (nt=1,682)	Low HIV-Prevalence $<$ 0.4% (nc=2,948)			
Socio-Economic Treatments (sample nt)	**Controls (sample size, nc)**	**CHINA (n =860)**	**MALAYSIA (n =261)**	**THAILAND (n =228)**
Urban Household (nt=2,866)	Rural Household (nc=1,764)	nt = 638 nc =222	nt = 147 nc = 114	nt = 138 nc = 90
Boys (nt=2,264)	Girls (nc=2,366)	nt = 429 nc = 431	nt = 129 nc = 132	nt = 114 nc = 114
Goes-to-School (nt=4,111)	Not Go-to-School (nc=519)	nt = 837 nc = 23	nt = 247 nc = 14	nt = 203 nc = 25
High Socio Economic Status (nt=785)	The other two SES (nc=3,845)	nt = 234 nc = 626	nt = 67 nc = 194	nt = 47 nc = 181
Medium Socio Economic Status (nt=1,701)	The other two SES (nc=2,929)	nt = 307 nc = 553	nt = 128 nc = 133	nt = 50 nc = 178
Low Socio Economic Status (2,144)	The other two SES (nc=2486)	nt = 319 nc = 541	nt = 50 nc = 211	nt = 131 nc = 97
Occupation of main income earner is UNSKILLED (nt=947)	Any other Skill level (nc=3,683)	nt = 82 nc = 778	nt = 20 nc = 241	nt = 55 nc = 173
Occupation of main income earner is MEDIUM-SKILLED (2,407)	Any other Skill level (nc=2,223)	nt = 553 nc = 307	nt = 126 nc = 135	nt = 128 nc = 100
Occupation of main income earner is SKILLED/PROFESIONAL (1,276)	Any other Skill level (Nc=3354)	nt = 225 nc = 635	nt = 115 nc = 146	nt = 45 nc = 183

In terms of the socio-economic treatments, urban, go-to-school status and gender are each binary effects. On the other hand, we can distinguish three SES levels, namely, high, medium and low. Because treatments between SES levels are not considered (e.g., the effect of high versus low SES), each SES related treatment is also defined as a binary outcome. Thus, the effect of high SES on P_A is such that the control group are children whose households relative income (weighted by city of residence and country) is below that of high income households. The variable Skill level is equally divisible in three categories: skilled or professional (e.g., doctors, teachers, businessman), medium skilled (e.g., blue collar workers), and unskilled (e.g., herder, trader, vendor or farmer).[39]

Table 3 shows a selection of 3 countries among the 17 possibilities. These countries are not selected at random. The selection shows countries from each of the three blocks but, at the same time, these countries have to be prominent with respect to some other characteristic. China is chosen to represent Communist style economies because it is a country whose administration is highly criticised for having neglected the importance of information campaigns. At the other extreme, Thailand has been praised for acting quickly on preventing its population, thus it is chosen to represent the Low Income countries. Malaysia is chosen the represent Western style economies because it shows the highest HIV-prevalence in the block, thus allowing for the effect of democracy on a country with a high HIV infection rate.

For each of the treatment effects in Table 3, the paper presents average treatment effects (ATE) and average treatment effects on the treated (ATET).[40] Estimates of ATET requires control variables that reflect characteristics associated with both the treatments and the outcome of interest P_A. Table 4 presents summary statistics for the selection of variables that might be associated with both treatment and outcome. The table presents statistics according to country blocks and for the three representative countries. The main ideas is that allowing for these variables the control sample is organized into cells, and each treated unit is compared to a counterfactual that has similar 'cells' characteristics.[41]

---

[39] This categories are difficult to decide. The aim of Skill level is to pick up some information on the possible knowledge of parents in the household. For example, a trader in Viet Nam might earn more income than a teacher in the same country. Thus, relative income would place the teacher in medium rather than high SES. But a teacher might be better at informing his or her child on the dangers of HIV/AIDS. This are the differences that one hopes to pick up by making a distinction between treatments referred to as SES and those with respect to Skill level. The are categories such as 'Government worker' where the survey does not distinguish between skill levels, or 'unemployed'. Thus, the treatments SKILL are viewed with some caution in this first version of the paper.

[40] See Section 4 for a detail explanation on these two measures of interest.

[41] To give an example, lets assume that we are interested on the effect of been a boy on the probability of avoiding infection, P_A, for those who are boys, relative to girls. If the covariates are 'urban' and 'go-to-school', the cells are 4 (GTS in U, GTS in R, Not GTS in U, and Not GTS in R). The outcome P_A of a treated unit (a boy) going to

Table 4: Sample Statistics (standard error and deviations) for a selection of covariates by country block and for three country representatives. Source: UNICEF 2001 Survey on Children's views in SEA and the Pacific.

By Country Block	Complete	Western	Communist	Low Income
Units	4,630	1,215	2,083	1,332
% Urban	61.9 (0.7)	82.6 (1.1)	62.7 (1.1)	41.8 (1.4)
% Boys	48.9 (0.7)	49.7 (1.4)	48.4 (1.1)	48.9 (1.4)
% Go to School	88.8 (0.5)	94.4 (0.7)	89.7 (0.7)	82.3 (1.1)
Man Age of selected [14-17]	15.5 (1.1)	15.5 (1.1)	15.5 (1.4)	15.5 (1.1)
% in 14-15 group	49.6 (0.7)	49.1 (1.4)	49.8 (1.1)	49.9 (1.4)
% in 16-17 group	50.4 (0.7)	50.9 (1.4)	50.2 (1.1)	50.2 (1.4)
Socio Economic Status				
% Low SES	16.9 (0.6)	16.9 (1.1)	18.1 (0.8)	15.2 (1.0)
% Medium SES	36.7 (0.7)	49.3 (1.4)	32.3 (1.0)	32.3 (1.3)
% High SES	46.3 (0.7)	33.7 (1.4)	49.7 (1.1)	52.5 (1.4)
Skill level main adult earner				
% Unskilled	20.5 (0.6)	7.9 (0.8)	24.1 (0.9)	26.2 (1.2)
% Semiskilled	52.0 (0.7)	47.1 (1.4)	54.9 (1.1)	51.9 (1.4)
% High Skill (professional)	27.6 (0.7)	45.1 (1.4)	21.0 (0.9)	21.9 (1.1)
% Get on well with my father	81.4 (0.6)	77.5 (1.2)	82.8 (0.8)	82.7 (1.0)
% Get on well with my mother	90.5 (0.4)	89.6 (0.9)	91.6 (0.6)	89.5 (0.8)
% Get on well with my guardian	4.2 (0.3)	3.1 (0.5)	2.5 (0.3)	8.0 (0.7)
% Expect better future relative to parent	83.4 (0.6)	81.2 (1.1)	83.2 (0.8)	85.8 (1.0)
% Expect better future conditions in community	75.4 (0.6)	70.7 (1.3)	76.4 (0.9)	78.0 (1.1)
% Experience Peer pressure	12.6 (0.5)	10.5 (0.9)	10.0 (0.7)	18.5 (1.1)
% Experience hunger	2.8 (0.2)	1.5 (0.4)	0.7 (0.2)	7.4 (0.7)
% Experience violence at home	14.0 (0.5)	15.7 (1.0)	14.5 (0.8)	11.8 (0.9)
Happiness Indicator: % Happy most of the time	48.5 (0.7)	55.6 (1.4)	54.6 (1.1)	32.4 (1.3)
% Happy Only sometimes	49.4 (0.7)	42.0 (1.4)	43.7 (1.1)	64.9 (1.3)
% Almost never happy	1.9 (0.2)	2.2 (0.4)	1.4 (0.3)	2.3 (0.4)
% Never experience happiness	0.3 (0.08)	0.2 (0.1)	0.3 (0.1)	0.4 (0.2)
By Selected Countries	**All three**	**Malaysia**	**China**	**Thailand**
Units	1,349	261	860	228
% Urban	68.4 (1.3)	56.2 (3.1)	74.2 (1.5)	60.5 (3.2)
% Boys	49.8 (1.4)	49.3 (3.1)	49.9 (1.7)	50.0 (3.3)
% Go to School	95.4 (0.6)	94.6 (1.4)	97.3 (0.6)	89.0 (2.1)
Man Age of selected [14-17]	15.5 (1.1)	15.6 (1.1)	15.4 (1.2)	15.3 (1.1)
% in 14-15 group	51.4 (1.4)	40.6 (3.0)	52.8 (1.7)	59.2 (3.3)
% in 16-17 group	48.5 (1.4)	59.4 (3.0)	47.2 (1.7)	40.8 (3.3)
Socio Economic Status				
% Low SES	25.8 (1.2)	25.6 (2.7)	27.2 (1.5)	20.6 (2.7)
% Medium SES	36.0 (1.3)	49.0 (3.1)	35.7 (1.6)	21.9 (2.8)
% High SES	38.3 (1.3)	25.3 (2.7)	37.1 (1.7)	57.5 (3.3)
Skill level main adult earner				
% Unskilled	11.6 (0.9)	7.7 (1.7)	9.5 (1.0)	24.1 (2.8)
% Semiskilled	59.8 (1.3)	48.3 (3.1)	64.3 (1.6)	56.1 (3.3)
% High Skill (professional)	28.5 (1.2)	44.1 (3.1)	26.2 (1.5)	19.7 (2.6)
% Get on well with my father	84.2 (1.0)	77.8 (2.6)	89.1 (1.1)	73.3 (2.9)
% Get on well with my mother	91.3 (0.8)	90.4 (1.8)	93.4 (0.9)	84.2 (2.4)
% Get on well with my guardian	4.3 (0.6)	3.8 (1.2)	3.2 (0.6)	9.7 (2.0)
% Expect better future relative to parent	91.8 (0.8)	87.4 (2.1)	93.1 (0.9)	91.7 (1.8)
% Expect better future conditions in community	89.3 (0.8)	85.1 (2.2)	90.7 (1.0)	88.6 (2.1)
% Experience Peer pressure	16.1 (1.0)	19.2 (2.4)	15.0 (1.2)	16.7 (2.5)
% Experience hunger	0.7 (0.2)	0.8 (0.5)	0.5 (0.2)	1.3 (0.8)
% Experience violence at home	12.3 (0.9)	19.5 (2.5)	8.5 (1.0)	18.4 (2.6)
Happiness Indicator: % Happy most of the time	69.5 (1.3)	49.4 (3.1)	86.5 (1.2)	28.1 (3.0)
% Happy Only sometimes	29.8 (1.3)	49.0 (3.1)	13.0 (1.2)	70.6 (3.0)
% Almost never happy	0.7 (0.2)	1.2 (0.7)	0.4 (0.2)	1.3 (0.8)
% Never experience happiness	0.07 (0.07)	0.0 (0.0)	0.1 (0.1)	0.0 (0.0)

school and living in a Rural area is compared to a control (a girl) in the cell GTS in R, since for a treated unit (a boy) with GTS in U, that cell has the best counterfactual individuals in the control group.

This selection of variables is an exhaustive list with respect to the information available in the survey, and, therefore, 'cells' formed with such variables provide the best comparative scenario between treated and control units. Some covariates in Table 4 define treatments in Table 3. For each case, these are used as covariates to form comparative cells if and only if they are not simultaneously treatment [42]

Table 4 shows that Western economies concentrate their populations in urban areas, while their teenagers are more likely to be at school, specially when compared to Low Income countries. SES distribution is similar in all three country-blocks, but the percentage of unskilled is significantly lower for those surveyed in Western economies than for the Communist or Low Income blocks. Peer pressure, hunger, violence at home and lower happiness index are all characteristics with greater probability for children in Low Income countries. The relative distribution of the variables among the three blocks is maintained when estimating similar distributions among the three representative countries.

Not all variables in Table 4 form comparative cells between treated and controls for each of the treatments considered, with the set of confounders varying across treatments, specially for treatments that deal with each of the three countries independently. The reason for this is that for some binary variables in Table 4 there is little variability between treated and control, and, therefore, they need to be taken out of the confounder set to allow estimation of the propensity score. All covariates in Table 4 enter the estimation process at levels.

4 Estimation Methods and Identification Conditions

The question this paper aims at answering is "*What is the effect of belonging to a particular social, economic or geopolitical group, for those children that belong to such particular social condition, on the probability that a child is sufficiently equipped to avoid HIV-infection, compared to the hypothetical state of belonging to a different (worse or better off) social group?*". This questions targets the causal relation from

[42] When acting as a covariate, the variable Skill enters as occupational category, that is takes 1 of 14 values.

political or socio-economic condition to outcome (i.e., P_A, the assigned relative probability of avoiding HIV infection given the child's knowledge, probability of risk taking behaviour and his or her country's specific HIV-prevalence). The econometric approach often used to answer questions of causal nature is given by Rubin's (1984) 'potential outcome' approach to causality.

The population of interest is the 2001 population of teenagers in the 17 surveyed countries (or territories) between the ages of 14 and 17, represented by the sample size n. Section 3 defines age as the only sample selection criteria. Each observed unit discloses information on knowledge of HIV/AIDS and proxies for risky-health-habits, while UNAIDS Fact Sheets provide comparable country's estimates of the HIV-prevalence.[43] These three pieces of information determine P_A as defined in Section 3.1, while Section 3.2 proposes independent geopolitical and socio-economic causal treatments, each independently levelled as 'T'. For example, on estimating the effect on P_A of belonging to a Communist style economy as opposed to a Western style economy, the treatment T is a binary outcome that equals 1 if the child's country of residence is in the Communist style block, and 0 if the child belongs to a Western style country.

Let T_i be the binary assignment indicator that determines if unit $i \in ns$ gets the treatment ($T_i = 1$) or not ($T_i = 0$), for $n \geq ns = nt + nc$, where nt and nc are the size of treated and control samples, respectively, and $n(\geq ns)$ is the complete sample. Let $P^1_{A,i}$ and $P^0_{A,i}$ be the potential assigned probabilities of not been able to avoid HIV-infection associated with the treated and untreated states, respectively. The notion "potential" is used to emphasis that only one of $(P^1_{A,i}, P^0_{A,i})$ is observed for every child in the sample. For any given treatment of interest T, $P^1_{A,i}$ is the 'actual observed' outcome ($P_{A,i}$) associated with an individual $i \in ns$ that complies with the geopolitical or socio-economic condition T. Likewise, $P^0_{A,i}$ is the 'actual observed' outcome ($P_{A,i}$) of a child in $ns(\leq n)$ with the counterfactual status defined by T.

[43] Except for East Timor. Information in this paper for independent territory of East Timor comes from The World Bank (http://devdata.worldbank.org/external)

Our parameter of interest, Δ^{TT}, is the mean effect on P_A of belonging to a particular country or socio-economic class (say Group J), rather than belonging to a (mutually exclusive) alternative, for those people that in fact belong to the Group J. The parameter is known in the literature as the 'Average Treatment Effect on the Treated', or ATET for short. In our context it can be expressed as follows:

$$\Delta^{TT} = E[P_A^1 - P_A^0 \mid T=1] = E[P_A^1 \mid T=1] - E[P_A^0 \mid T=1] \tag{1}$$

Clearly, Δ^{TT} is not identified by the data since identification of the causal effect would require observing the counterfactual outcome to $P_{A,i}^1$ (i.e., $P_{A,i}^0$) for each i^{th} unit in the treated sub-sample, thus allowing us to estimate $E[P_{A,i}^0 \mid T=1]$. The problem would be solved if we assume randomness between treatment T and outcome P_A, that is, $E[P_A^0 \mid T=0] = E[P_A^0 \mid T=1]$. With this, we could use the average outcome of the control-sample as the average counterfactual outcome for the treated units to estimate (1). An example is sufficient to show why this might lead to a wrong policy conclusion. Let $T = urban$ for a particular country, that is, we want to understand the effect on P_A of belonging to a urban household, on urban children, relative to the alternative status of belonging to a rural household. Lets assume that , allowing for sampling error, sample-analogue estimates show that $(1/nt)\sum_{urban} P_A$ is significantly smaller than $(1/nc)\sum_{rural} P_A$, that is, the (average) treatment effect of living in a metropolis puts children at a much better position to avoid HIV-infection in the future. However, it could be that the number of sampled children in urban areas attending school is disproportional to that of surveyed children in rural areas, and that differentials in sample averages is picking up the schooling effect on knowledge of HIV/AIDS and health habit formation. Thus, comparing sample averages and neglecting going-to-school status might lead policy makers to wrongly conclude that information campaigns are failing children in rural areas, when in fact information campaigns do reach rural areas but only for children who attend school, and, therefore, it is schooling status that needs to be targeted.

In general, Section 2 shows sufficient evidence to suggest that income inequality, future expectations, education or even peer pressure are highly correlated to distinct social classes between and within countries, while these are also determinant factors on the evolution of HIV/AIDS. Section 3.2 suggest a set of X variables that might determine both selection into social class and HIV-status. Then, identification of Δ^{TT} is possible since conditioning on X implies that within sub-groups (e.g., those going to school, those not going to school), been untreated (i.e., living in a rural setting) is unrelated to what the outcome P_A would have been if the child had been treated (i.e., if the child lived in a urban area). This assumption is formalized in the econometrics (treatment effect) literature as The Conditional Independence Assumption (or CIA), and is given by:[44]

$$P_A^0 \perp T \mid X, \quad \forall x \in c, \quad c \subseteq \mathbb{R}^p \tag{2}$$

Therefore, $E[P_A^0 \mid T=1, X] = E[P_A^0 \mid T=0, X]$, and Δ^{TT} is identified such that,

$$\Delta^{TT} = E[P_A^1 \mid T=1] - \underset{X|T=1}{E}[E[P_A^0 \mid T=0, X=x] \mid T=1] \tag{3}$$

The CIA is a workable assumption as long as it holds for all variables in X, but does not account for unobserved characteristics that might also play a role in selection (e.g., cognitive ability of the child to pick up information, low discount rate of the future, parental motivation, etc).[45]

---

[44] The methodology employs a weak CIA. A stronger version than (2) is given by $(P_A^1, P_A^0) \perp T \mid X$. It would imply, for example, that going to school is truly exogenous with respect to both urban and rural status. In the present case, all treatments considered (e.g., geopolitical, gender, school-status, urban) can be successfully argued as truly exogenous with respect to the outcome P_A, since a child's knowledge and behaviour is are posterior to any of the treatments for any given child. This, together with the weak CIA assumption in (2) for the sub-population of untreated, are sufficient identifying assumption for the ATET.

[45] Black and Smith (2003) motivate the argument that the CIA is nothing but an assumption, as could be the assumption of a common support condition in the regression analysis. Where unobserved factors play an important role in selection between treatment, the CIA does nothing to control for it. The believe, however, is that sufficient information in the dataset, as is the case here, helps to overcome such limitations.

Δ^{TT} can be estimated using the sample analogue, provided that for every treated unit there is a comparison unit in the untreated sub-group with similar X characteristics (e.g., if the sub-sample of children in urban households go to school, the rural household sub-sample must also contain children going to school). This is known as the common support condition, and in our case it can be expressed as $P(T=1|X=x)<1$ for all X. The implication is that there is a common overlap between the distribution of X's in the two states defined by the treatment. This assumption is the cornerstone for matching methods as these rely on finding the best 'twin-observation' among the untreated units, for each unit in the treated sub-sample. Furthermore, controlling for possible common support problems is an advantage over standard regression analyses. This is because where no positive density of X exists between treated and untreated, regression analyses still projects the counterfactual outcome with reference to the functional form, even if this is a misspecification with respect to the true outcome equation. On the other hand, matching allows for checks on common support problems and provides alternative methods to deal with situations where the common support might not be well defined by the data.[46]

When X is of high dimension, estimation of $E[P_A^0 \,|T=0, X=x]$ using nonparametric techniques, such as Kernel weighting, is subject to the curse of dimensionality (i.e., very low density per cell) and, therefore, increases imprecision in the parameter's estimate, specially around the tails of the distribution. Rosenbaum and Rubin (1983) show that it is sufficient to compare observations having the same conditional treatment probability, $P(T=1|X=x)=p(x)$, where $p(x)$ is also known as 'the propensity score'.[47]

[46] See Black and Smith (2003) for an application of different checks on the common support problem, and the effect on the ATET when accounting for possible failures on the common support.

[47] The method amounts to weighting variables in X according to their effect on the treatment, thus adjusting for difference in the confounders between treated and controls. In practice, if the treatment is binary, the ideas is to estimate an index model where the outcome is 1 for the treated units. The coefficients of X are used to predict the probability of treatment for each of the two sub-samples, where such predictions are the estimated propensity scores. For example, assuming a Probit model, the propensity score for treated and controls are given by $\Phi(x_t'\hat{b})$ and $\Phi(x_c'\hat{b})$, where $\hat{b}$ is the vector of estimated coefficient in $P(T=1|X=x;b)$.

Conditioning on $p(x)$ rather than X reduces the estimation problem to a one dimension and the estimate of Δ^{TT} is based on:

$$\Delta^{TT} = E[P_A^1 \mid T=1] - \underset{p(x)|T=1}{E}[E[P_A^0 \mid T=0, p(x)] \mid T=1] \quad (4)$$

The RHS of expression (4) has two components. The first component is the mean value of the outcome in the treated population. The second components it the mean value of the counterfactual population, where each counterfactual outcome results from matching on the propensity score. Matching methods differ between them according to how they weight each observation in the control (or comparison) group. For example, matching (on the propensity score) with the 'Nearest Neighbour' implies placing the full weight in one observation, that is, the propensity score for each i^{th} unit in the treated group is compared to all point estimates of the propensity score in the untreated group, and, for each $i \in nt$, the assigned counterfactual outcome $P^c_{A,j \in nc}$ is the outcome of the j^{th} control unit that minimises $| p_i(x_t) - p(x_c)|$. An alternative matching method is the use of a Kernel weight, for example, a Gaussian Kernel. This alternative differs from the Nearest Neighbour in that a Kernel function can assigns a non-zero weight to more than one (if not all) observations in the control sub-sample. For each i^{th} observation in the treated sub-sample, the function of a Kernel is to assign a weight w_j to each j^{th} control observation. The assigned weight w_j is determined by the distance between $p_i(x_t)$ and $p_j(x_c)$, relative to the distance between $p_i(x_t)$ and a sub-group $p_g(x_c)$, $g \neq j$ where $p_g(x_c) \subseteq p(x_c)$. The sub-group is defined by a bandwidth that breaks down the domain of the propensity score into blocks (or windows) such that comparing distances between the two vectors of the propensity scores becomes more efficient.[48] For each i^{th} treated observation, the estimated vector of j Kernel weights $w^i = (w^i_1, \ldots, w^i_{j-1}, w^i_J)'$, $j \in nc$, sums to 1. The counterfactual outcome to each $i \in nt$ is a weighted sum of the control's outcomes given by $P^c_{A,i} = \sum_{j=1}^{J} w^i_j P^c_{A,j}$. No weighting methods can be thought

[48] For example, Gaussian Kernel consists on estimating many local Gaussian densities, with points too far away from the mean value of the local density been given a zero weight, while those points that are close to the local mean define the window's sub-group.

as superior than others when matching on the propensity score, specially if the common support condition is met by the data.[49] However, with large control groups the use of Kernel weights can significantly reduce the variance on the estimates of Δ^{TT}. This paper employs Gaussian Kernels to weight the propensity score when matching between control and treated sub-samples, with the choice of bandwidth based on Silverman (1986) normal approximation.

As defined in Section 3.1, a simulation process assigns 200 values of P_A for each child in the sample of 4,630, thus allowing a distribution of possible assigned probabilities of avoiding HIV-infection to each surveyed unit. P_A is independent of the treatment under study, and drawn from a Beta distribution $B(a,b)$ with the centrality parameters determined by the child's knowledge on HIV/AIDS, his or her attitude towards risky health-habits and the child's country specific HIV-prevalence. With 200 vectors of the outcome variable P_A, we end up with 200 estimates of Δ^{TT} for each of the treatments under study. In each case, the average of these 200 estimates is the final outcome. Likewise, the sampling variance on each of the 200 estimates of Δ^{TT} is obtained using a naïve bootstrap procedure that consists on re-sampling 300 times (with replacement) from the original sample (as defined by each treatment). This bootstrap procedure assumes that each treatment-designed sample is representative of children age 14 to 17 in the population.

5 Results

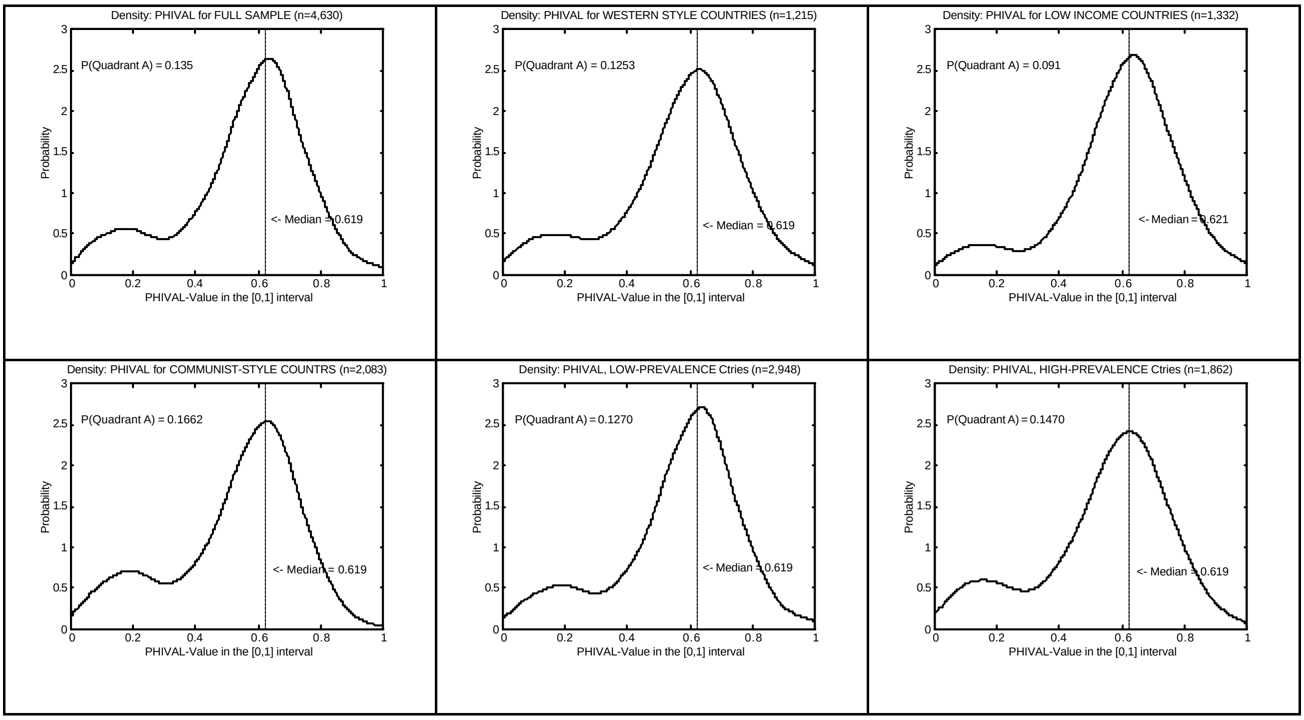
This section summarizes the results of testing multiple treatments on P_A differentials. The first round of estimation is concerned with differentials between country blocks. The second round is concerned with P_A differentials by socio-economic status both within-country and with respect to the complete sample. These estimates are all based on matching on the propensity score, with Appendix D complementing both sub-sections with tabulated results.

[49] See Lechner and Vazquez-Alvarez (2003) for a detail discussion and empirical application on the different choices of matching methods, alternative Kernel functions and alternative choice of bandwidth parameters.

5.1 Geopolitical Differentials on P_A

Nobel laureate and development theorist Amartya Sen has suggested that lack of democracy is a factor as important as poverty in propagating HIV/AIDS, mainly because non-democratic regimes fail at being accountable to their populations, while their countries often lack the appropriate infrastructure to effectively disseminate information (Sen (2002), Manning (2002), Patterson (2000)). If this was the case, one would expect to find that children in Communist oriented regimes (China, Cambodia, PDR Laos, Mongolia, Myanmar and Viet Nam) are less able to cope with the HIV/AIDS situation in the region, relative to their counterparts in other country blocks (either High or Low Income free democracies). That is, the estimated P_A density for surveyed children in the Communist block would be to the right of any other block definition (i.e., higher density at values of P_A close to 1). Table 5 displays the smooth density estimates by block definition: Appendix D complements these by plotting the density for each of the 17 countries (or territories) in the survey.

Table 5: Unconditional density estimates of P_A by Country Blocks (mnemonic PHIVAL = P_A)



All country blocks have estimated P_A-densities with similar shapes and medians, each mimicking that of the overall sample (upper left-hand cell in Table 5). However, the probability of finding children in Communist style regimes inside quadrant A (i.e., $P_A \leq 0.30$),[50] is significantly higher than for any other type of country. That is, in Communist countries, P(Quadrant A)=0.1662, while for Western style P(Quadrant A)=0.1253. Allowing for sample size ($n_{communist} = 2{,}083$, $n_{west} = 1{,}215$) , the null of 'equal density of children in Quadrant A' is rejected with a 0.07% significance level (i.e., $\hat{z} - test$ =3.27). A test of Communist against Low Income countries ($n_{lic} = 1{,}332$) results on $\hat{z} - test$ =6.63 so that the null of equal density in quadrant A between the two blocks is rejected with a p-value of 0, suggesting that children in Communist regimes display a distribution of P_A that makes them much better off at dealing with the HIV/AIDS epidemic, as compared to children in Low Income countries. Another conclusion from Table 5 is that awareness of HIV/AIDS is also significantly higher for countries with a higher (rather than lower) HIV-prevalence, where I have decided to place the cut off point between high and low at a prevalence of 0.4%.[51] This might indicate that children in countries with older epidemics are more knowledgeable, and display lower risk behaviour, for the simple reason that having people in the neighbourhood dying of the virus is in itself informative, irrespective of information campaigns. Although children in the Communist block seem better equipped, overall, the clear piece of evidence from Table 5 is that allowing for all 17 States, the vast majority of children are badly equipped to deal with the HIV/AIDS problem faced by the region, since only 13.5% of these have a ranking of P_A that places them on a 'safe' quadrant with respect to avoiding HIV infection, while more than 50% show P_A values associated with poor or no knowledge on HIV/AIDS and high-risk health behaviour (Table 5, Plot 1).

---

[50] The Quadrant A, B, C or D is determined by the combinations (k, r) as in Table 2. Appendix C, Table C2a shows that Quadrant A draws values of P_A from Betas with a mean below or equal to 0.30, Quadrant B draws from Betas with means between 0.30 and 0.45, Quadrant C from Betas with means between 0.50 and 0.65, and Quadrant D draws from Betas with means between 0.65 and 0.90. In all cases, Table C2b shows how each child's country HIV-prevalence adds a positive shock to each P_A draw.

[51] See Table 1, Section 3.1 for a distribution of the countries between these two categories.

So far the results are based on comparing densities. These do not control for P_A-differentials that might be driven by the socio-economic distribution of children in different country blocks, meaning that differentials do not necessarily reflect geopolitical differences. For example, children in the Communist block might be from urban areas that go to school, both factors affecting their knowledge and risk behaviour. If these two effects drive the results, we would not be able to conclude that 'Communism' in itself is the cause for these children to be better prepared at avoiding HIV infection. To control for these and isolate the effect of a treatment on P_A, any estimates discussed from this point onwards are based on matching methods. This implies that a treatment (e.g., been in a Communist country) is contrasted against the alternative (e.g., Westernised economies), by comparing treated children with their counterfactual, these latter chosen according to similar socio-economic characteristics. Any treatment considered is binary in nature (see Table 3, Section 3.2), thus, as explained in Section 4, matching on the propensity score is performed by means of Probit estimates. Then, for each treatment considered, $\hat{P}(T=1|X_t)=\Phi(x_t'b_t)$ and $\hat{P}(T=1|X_c)=\Phi(x_c'b_t)$, the propensity scores for the treated and untreated, respectively, have to show overlapping density estimates. Appendix D shows this to be the case for all treatments, where X is a vector of covariates that includes most variables as defined in Table 4 (Section 3.2). The implication is that the comparison between treated and their counterfactual is well defined with respect to common support conditions. The choice of matching method is based on Kernel weights with a normal bandwidth approximation.

We consider the effect of 7 different geopolitical treatments as defined by treatments T1 to T7 in Table D3, Appendix D.[52] Summarizing the effects we look at the average mean of the value P_A, and how this average changes with respect to the counterfactual for any of the treatments.

Testing for P_A differentials between High Income (Western style) democracies (Australia, Hong Kong, Macao, Singapore, Malaysia, South Korea) versus all other countries shows that Westernised children

[52] Table D2 shows results of ATE, therefore based on the assumption that socio-economic background attributes are not determinants on the differential of knowledge and risk behaviour.

display an average of P_A to the right of their counterfactuals; the estimate shows that, relative to other children in the region, they are 5% worse off equipped to deal with the HIV/AIDS problem. The same conclusion is applicable when comparing Low Income countries to all other countries: the ATET estimate suggests that children in Low Income economies are 7% worse off. The opposite is true for children in Communist style economies since, relative to their counterfactuals in other economies, their average P_A value of 0.5285 would shift up to 0.5777 were they to be living in other types of economies. This means that if they lived in a country block other than in the Communist one to which they belong, their P_A distribution would shift to the right by 0.05 points (in the [0,1] interval), making them 9.3% worst off in terms of knowledge and risk behaviour. Although this seems small, given the small variability in the difference between $P_{A,communist}$ and $P_{A,non-communist}$, such differential is highly significant with a 95% confidence interval on the ATET well away from crossing the zero point (see Table D3, Geopolitical treatment T3). Since all socio-economic confounders are accounted for, we would interpret the difference as driven by the positive effect that information campaigns have on children in Communist economies, relative to the effect that such campaigns might have on children of other countries. In principle this result would seem to contradict Sen's theoretical implication of non-democracies on HIV/AIDS knowledge and risk behaviour. To make sure the result is not driven by the very bad P_A distribution of a few countries (see these distributions in Appendix D), we select two countries, Thailand and Singapore, and compare each of these to each Communist country, independently. The choice is not random. Thailand was one of the first countries to use preventive information campaigns to reduced the HIV problem in the early 1990s, with budgets to the tune of USA$35 million per programme. Comparing Thailand versus each Communist country is a check on the real understanding that 'Communist' children may have on the HIV/AIDS epidemic, as compared to children in countries with well organized longitudinal information campaigns that should make their populations relatively well informed (see Density plot of Thailand in Appendix D). Singapore also fairs very well in terms of P_A with its children showing to be better informed and better equipped to deal with the epidemic, relative to other High Income Western style economies in the survey. But Singapore has a relatively low HIV-prevalence when compared to Thailand, meaning that comparing Singapore, rather than just Thailand, to each Communist country helps us to control for the possibility that a high prevalence might be informative in itself.

Table 6 shows that China, PDR Laos, Mongolia, and Myanmar are each significantly worse off than Thailand: in each case, the P_A distribution for Thai counterfactuals is to the left of the actual P_A distribution of either Chinese, Laos, Mongolian and Myanmar children. On the other hand, relative to Thailand, there is no significant P_A differentials with respect to Cambodia, while differentials with respect to Viet Nam shows that Vietnamese children are much better off than their Thai counterparts in dealing with the HIV/AIDS epidemic, in fact, 22% better off.

Table 6: **Average Treatment Effects on the Treated (ATET), Treatments (Standard errors in brackets) for each communist country independently against two selected control countries (Singapore and Thailand).*(000) significant at 10%, 5.0%***

Treatment Code	Treated → Control	$E[P_A^T \mid T]$ (a)	$E[P_A^C \mid T]$ (b)	Δ^{TT} (a)-(b)	95% Confidence Intervals
T-CHINA	CHINA → SING	0.5457 (0.0062)	0.4762 (0.0232)	***+0.0696 (0.0341)***	***(+0.0279, +0.1083)***
	CHINA → THAI	0.5457 (0.0065)	0.5094 (0.0231)	***+0.0363 (0.0239)***	***(-0.0300, +0.1116)***
T-CAMBODIA	CAMB → SING	0.4769 (0.0164)	0.5259 (0.0278)	-0.0490 (0.0333)	(-0.0955, -0.0050)
	CAMB → THAI	0.4769 (0.0154)	0.4900 (0.0293)	-0.0131 (0.0323)	(-0.0491,+-0.0262)
T-LAOS	LAOS → SING	0.5445 (0.0117)	0.5229 (0.0230)	+0.0217 (0.0255)	(+0.0010, +0.0431)
	LAOS → THAI	0.5445 (0.0129)	0.4874 (0.0230)	***+0.0571 (0.0267)***	***(+0.0087, +0.1012)***
T-MONGOLIA	MONG → SING	0.5615 (0.0111)	0.4905 (0.0249)	***+0.0710 (0.0267)***	***(+0.0288, +0.1110***
	MONG → THAI	0.5615 (0.0103)	0.4980 (0.0239)	***+0.0635 (0.0267)***	***(-0.0183,+-0.1439)***
T-MYANMAR	MYNA → SING	0.6912 (0.0170)	0.5634 (0.0354)	***+0.0555 (0.0240)***	***(+0.0323, +0.0822***
	MYAN → THAI	0.6912 (0.0161)	0.5273 (0.0273)	***+0.0797 (0.0238)***	***(+0.0112, +0.1438)***
T-VIET NAM	VIET → SING	0.3935 (0.0154)	0.5643 (0.0247)	***-0.0865 (0.0294)***	***(-0.1261, -0.0452***
	VIET → THAI	0.3935 (0.0143)	0.5308 (0.0297)	***-0.0803 (0.0244)***	***(-0.1316, -0.0286)***

Clearly, compared to a country that has put a huge amount of resources on preventive campaigns (Thailand), children in Communist countries are doing much worse off at dealing with the HIV/AIDS epidemic. The only two countries that do either as well or better than Thailand (Cambodia and Viet Nam) can be explained. Viet Nam has achieved remarkable economic progress over the last 10 years, opening up as a new emerging capitalist oriented economy in the region, and since 1999 has become actively involved on ASEAN, a Task force on AIDS addressing HIV/AIDS related issues at sub-regional level (UNAIDS, 2002).

This, together with the fact that it is one of the new emerging focus of the epidemic might explain why children are more informed. Cambodia, on the other hand, is a country destroyed by war with poor means to disseminate information and yet, differentials on P_A between Cambodian and Thai children are insignificant: this might be explained by the fact that Cambodia borders with Thailand from the West and with Viet Nam from the East, and regional proximity might imply that information campaigns have positive external effects on neighbouring countries. Another reason for the high knowledge on HIV/AIDS and low risky health behaviour of children in Cambodia could be due to the country's high HIV-prevalence: with an infection rate that translates onto 3 in every 100 adults being HIV-positive it might not be surprising that these children are aware of someone who is directed or indirectly affected by AIDS, thus, the epidemic in itself is the main source of knowledge. In fact, when compared to Singapore, a country that has a low HIV prevalence, Cambodian children shows a mean P_A value that makes them 10.3% better off than their Singaporean counterparts. Due to the huge variability in P_A differentials between the two countries, this estimate is not significant, not even at a 1% level. Table 6 shows that any country in the Communist block, other than Cambodia, are significantly worse off equipped than Singapore to deal with the HIV/AIDS epidemic.

If border proximity to both Thailand and Viet Nam can explain the high level of knowledge in Cambodian children, the puzzle is why is there such remarkable differential between Thailand and PDR. Laos or Myanmar, since both countries also share borders with Thailand. In terms of PDR Laos, the data cannot be used to test for differentials on P_A by border proximity to Thailand or Cambodia, since all children surveyed in PDR Laos come from the Capital Vientiane or its surroundings. However, the 262 children surveyed in Myanmar can be distributed between closeness to the Thai border (Yangon sub-urban, Pathein region and Mundon region), or far from the Thai border (Kawal near India, and Phatheigyi near India and China). Testing for the ATET of being a child in Myanmar close to the Thai border, relative to a Thai counterpart shows that the mean value of P_A for such Myanmar children is 0.5661. This improves only slightly the value of 0.5730 for the complete sample in Myanmar. The Thai counterparts to Myanmar's children in the Thai border would end up with a mean value of P_A equal to 0.4960. Thus, the ATET is 0.0710 and positive. This means that

Myanmar children in the Thai border are 12.5% worse off than their Thai counterparts. Allowing for sampling error the estimates are significant. Thus, it seems that border proximity to a (in theory) knowledgeable country such as Thailand is not a factor affecting Myanmar's children stock of knowledge.

A worry is that the above results are driven, not by a combination of HIV/AIDS knowledge and risky health behaviour, but by country's differentials in risky health behaviour only. An interesting exercise is to see what happens to estimates of ATET if P_A was constructed using only HIV/AIDS knowledge. Appendix E shows this, with the resulting new P_{A2} such that $P_{A2} = B(a, b | k) + B(a, b | HIV_{country})$. Tables E4 and E5, show that the results in terms of country block comparison are not qualitatively different than those attained with P_A. However, when comparing each communist country independently to both Singapore and Thailand there are some significant variations in the results. Using P_{A2} Cambodian children are now significantly better off than Singaporeans, whereas with P_A the difference was not significant. This shows that relative to Cambodia, those in Singapore are less likely to engage on high health risk behaviour, but, for any given level of risk behaviour, Cambodian children are more knowledgeable on the dynamics of HIV/AIDS. The second change is that with P_{A2} PDR Laos becomes significantly worse off than Singapore. Thus, relative to Singapore, children in Laos are worse off in terms of knowledge irrespective of low health risk behaviour. Comparing results for Viet Nam between Tables 6 and Table E5 (Appendix E) shows that not accounting for risk reinforces the ATET results for Viet Nam. This means that if knowledge was the only factor, Vietnamese children would be 43.4% and 34.9% better off at coping with HIV/AIDS than Singaporean and Thai children, respectively. Accounting for risky behaviour, these percentages would drop to 20.4% and 18.9%, also respectively. This seems to suggest that although children in Viet Nam might show greater chances of becoming engaged on high risk behaviour, relative to their Thai and Singaporean counterparts, either with or without risk, on average, Viet Nam children are better equipped to deal with the problems posed by HIV/AIDS. Other than Viet Nam and Cambodia, evidence would support Sen's claim that non-democratic regimes are worse off at equipping its population to deal with the HIV/AIDS epidemic, since children in 'most' communist countries display knowledge and risk behaviour that puts them at a critically worst position to avoid infection, relative to countries with non-communist type of regimes.

5.2 Socio-Economic Differentials on P_A

Instead of country block differences, the effects are now with respect to a selection of socio-economic indicators, namely urban households (versus rural), been a boy (as opposed to a female child), and going to school (as opposed to not attending school). We also test for differentials with respect to SES status (high, medium and low) where SES is defined by relative household income, and for occupational skill differentials (i.e., main income earner in household is either unskilled, medium skill or professional).

Results in Table D3, Appendix D, show that using the complete sample, living in a urban area makes children significantly better off (by 3%) relative to children in rural settings, been a boy significantly reduces their position to deal with HIV/AIDS relative to girls (by 3%), while the strongest effect is, not surprisingly, that of going to school: those who attend school are 12.9% better off than those who do not go to school, controlling for any other socio-economic condition. Household income (the underlying identification of SES status) has no significant effect on the child's P_A value-position between different SES groups. Three countries are selected, one for each of the country blocks, to estimate the same set of socio-economic effects. The results suggest that in China, belonging to a urban household has a positive effect on the children's capacity to deal with HIV/AIDS, while the opposite is true for Malaysia. The country representing Low Income economies (Thailand) shows that urban/rural position is not a significant factor determining differentials on P_A. However, for Thailand, going to school is a significant factor, with Thai children at school 21.4% better off than their counterfactuals outside the school system. In both China and Malaysia estimates of the ATET on P_A makes children going to school better off at dealing with the HIV/AIDS problem, relative to those who do not attend school, but in both countries the estimates are not significant.

6 Conclusions

This paper aims at evaluating geopolitical and socio-economic effects on differentials in the probability of becoming HIV-positive. The assumption is that successful information campaigns should disseminate the correct knowledge on how to avoid HIV/AIDS to all sectors in the population, and if so, differentials in the capacity to avoid HIV infection – as given by knowledge and risky health behaviour – should not differ significantly between socio-economic groups. Likewise, finding geopolitical differences could indicate information campaigns' inequality between countries in a given region, and, therefore, country's inequality in their ability to control the HIV/AIDS epidemic.

The empirical analysis draws from a survey of children representing 17 countries (and territories) in South East Asia and the Pacific. The survey was organized and collected by and for UNICEF during the first half of 2001. The selected sample comprises 4,630 children aged between 14 and 17 at the time of the interview. These are children that today, in 2004, have already entered the adult population, while, at the time of the survey (at least) 4% already indicated to be engaged on HIV risk-related behaviour. All sampled children were born between 1983 and 1987, meaning they had already been young adolescents for at least 2 to 5 years at the time of the survey, and, therefore, sufficiently old to be aware of their social surrounding. The survey collects information on each child's precise knowledge on HIV/AIDS as well as indicators with respect to risky health behaviour. The paper assumes that both their knowledge on HIV/AIDS and attitude towards health risk contain the effect of regional information campaigns on HIV/AIDS and related issues (e.g., other STD, Malaria, Hygiene, etc).

Based on declared HIV/AIDS knowledge, on the probability of risky health behaviour and allowing for country specific random shocks, the paper proposes the construction of P_A, the within-sample relative probability that a child is able to avoid infection once he or she becomes an adult and starts to face HIV risk situations. The construction of P_A assumes that knowledge, risky health attitudes and a country's HIV-prevalence are the three major factors affecting the chance of infection. This assumption is based on evidence that describes the evolution of the HIV/AIDS epidemic over the past 20 years.

Statistical analyses of the outcome variable P_A suggest that, overall, children in South East Asia and The Pacific Region are badly equipped to deal with situations that might lead them to become HIV positive. Only 13.5% show a combination of HIV/AIDS knowledge and risk behaviour that would place them on a 'safe' position in terms of avoiding infection, while more than half in the sample have very poor knowledge (if any at all) and are likely to be engaged on risky health behaviour (e.g., alcohol drinking, sniffing glue or the use of illicit drugs). A policy recommendation based on these estimates is that both national and international agencies in South East Asia and the Pacific ought to design information campaigns that significantly increase the capacity of children to deal with the HIV/AIDS epidemic. UNAIDS Fact Sheets suggest that all surveyed countries have national strategies aimed at increasing awareness in the population. The empirical section based on UNICEF's 2001 survey shows that these campaigns are not successful at targeting children to modify their behaviour before they enter adulthood, something that should worry authorities in a region classified as the second most affected in the world, holding some 20% of the global HIV-positive population and with clear emerging focuses of the epidemic.

The outcome P_A is compared between countries according to specific country block definitions, namely, Western style economies, Low Income economies and Communist style economies. The comparison is performed using matching methods to allow for the counterfactual approach associated with treatment effect techniques for programme evaluation. This approach means that comparisons are performed while controlling for observed socio-economic differences, so that finding P_A-differentials can be attributed to specific geopolitical treatments. On a first round of estimation the findings suggest that relative to any other block, but also with respect to both Western and Low Income countries, children in the Communist block are significantly better equipped to deal with an emerging HIV/AIDS epidemic. This initial result would contradict the prediction by Nobel laureate and development theorist Amartya Sen. His theoretical arguments suggest that 'lack of democracy' is a factor as important as poverty in propagating HIV/AIDS both because non-democratic regimes are not accountable to their populations, and because it is often the case that the infrastructure in Communist countries is insufficient for the effective dissemination of information (Sen (2002), Manning (2002), Patterson (2000)). To examine this further, the results section contrasts each communist country (independently) against both Thailand and Singapore. Thailand is chosen because it is a country that has used active preventive campaigns to stop the virus taking over the full population. Singapore is selected because it has also done well at informing its population, but, at the same time, the HIV-prevalence is low relative to that of Thailand's. Compared to these two, most Communist countries show children that are significantly worse off at dealing with the HIV/AIDS epidemic. For example, Chinese children are 12.7% and 10% worse off than Singapore and Thailand, respectively, at coping with the HIV/AIDS problem. The only two countries classified as Communist whose children seem better equipped to deal with HIV/AIDS are Cambodia and Viet Nam. Viet Nam might be explained because it is opening towards a more free social system that has allowed the intervention of preventive campaigns by international organizations. Cambodia is more difficult to explain. Perhaps its old epidemic has lead to such high HIV-prevalence that the epidemic in itself is the source of knowledge to the new generations of children. Any of the results are qualitatively the same if we control or not for possible differences due to differentials in the probability of high risky health behaviour between countries.

In terms of socio-economic treatments, the results suggest that living in urban households, being a girl, and going to school are all significant factors that imply a better position with respect to a child's capacity to deal with the HIV/AIDS epidemic. Outstanding is the result that going to school makes children 12.9% better off at dealing with a situation of high HIV risk, a result that increases to 21.6% in countries such as Thailand. Based on these results, there are clear policy implications in that information campaigns aimed at targeting children should also be design to reach those who are outside the school system, specially if they live in a non-metropolitan area. If this policy interpretation is ignored the epidemic could perpetuate itself among those who do not have access to information campaigns due to their social or economic condition, even in the presence of expensive (country or regional) interventions.

References

Angrist, J. (1991), "Instrumental variable estimation of average treatment effects in econometrics and epidemiology", *NBER WP Series*, No.115.

Black, D., and J. Smith (2002), "How robust is the evidence on the effects of college quality? Evidence from matching", *University of Maryland working paper*.

Cornia, G.A. (2002), "Overview of the impact and best practice responses in favour of children in a world affected by HIV/AIDS", in *AIDS, Public Policy and Child Well-being*,(ed. G.A. Cornia), UNICEF Innocenti Research Centre.

Cornia, G.A, and F. Zagonari (2002), "The HIV impact on the rural and urban economy", in *AIDS, Public Policy and Child Well-being*,(ed. G.A. Cornia), UNICEF Innocenti Research Centre.

De Walque, D. (2001), "How does the impact of an HIV/AIDS information campaign vary with educational achievement? Evidence from Rural Uganda", *Department of Economics WP Series*, University of Chicago.

Dow, W.H and T. Philipson (1996), "An empirical examination of the implications of assortative matching on the incidence of HIV", *Journal of Health Economics*, 15(1996), 735-749.

Global AIDS Policy Coalition (Updates, 1996, 2002), *AIDS in the World Series* (Harvard Press, Cambridge).

Heckman, J. J., H. Ichimura, and P. Todd (1997), "Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme", *Review of Economics Studies*, 64, 605-654.

Heckman, J. J., and J. Hotz (1989), "Choosing among alternative methods for estimating the impact of social programs: The case of manpower training", *Journal of the American Statistical Association*, 84, 862-874.

Janjaroien, W.S, and S. Khamman (2002), "Perinatal AIDS mortality and orphanhood in the aftermath of the successful control of HIV epidemics: The case of Thailand", in *AIDS, Public Policy and Child Well-being*,(ed. G.A. Cornia), UNICEF Innocenti Research Centre.

Lechner, M. (1999), "Continuous off-the job training in East Germany after unification", Journal of Business and Economics Statistics, 17, 74-90.

Lechner, M. and R. Vazquez-Alvarez, (2003), "The effects of disability on labour market outcomes in Germany: A matching approach", *HSG WP0320, University of St.Gallen*, Switzerland.

Manning, R. (2002), "AIDS and democracy: What do we know?", *Research Intern Health Economics and HIV/AIDS Research Division (HEARD), University of Natal, Durban* (for the proceedings of the Cape Town Workshop on AIDS, HEARD, DARU, IDASA).

Patterson, D. (2000), "Political commitment, government and AIDS: A discussion Paper", *Interagency Coalition on AIDS and Development*, Dec.1 2000 communicate.

Pegatienan, J. and D. Blibolo (2002), “HIV/AIDS, lagging policy response and impact on children: The case of Côte d’Ivoire”, in *AIDS, Public Policy and Child Well-being*,(ed. G.A. Cornia), UNICEF Innocenti Research Centre.

Rosenbaum, P., and D. Rubin (1983), “The central role of the propensity score in observational studies for causal effects”, *Biometrika*, 70(1), 41-55.

Rubin, D. (1974), “Estimating casual effects of treatments in randomised and non-randomised studies”, *Journal of Educational Psychology*, 66, 688-701.

Sen, A. (2002), “How to judge Globalism”, *The American Prospect*, vol.13, Jan.2002.

Silverman, R. (1986), “Density estimation”, *London: Chapman and Hall*.

Smith, J. and P. Todd (2002), “Does matching overcome LaLonde’s critique of non-experimental methods?”, *Journal of Econometrics*, forthcoming.

The Economist (December, 2003), “Help at last (Leading article, pg. 13)”, “AIDS: A mixed prognosis (Science & Technology section, pg.87-89)”, Vol.369,No.8352.

The Economist (January, 2004), “A Special Economist survey on Sub-Saharan Africa”, Vol.370, No.8358, 1-16.

The Guardian Newspaper (9 October, 2003), “Vatican: Condoms don’t stop AIDS!”, *comments on BBC1 Panorama, October 2003*, by Steven Bradshaw.

The Guardian Newspaper (February 6, 2003), “Tests show no threat of imported epidemic”, by James Meikle.

The Guardian Newspaper (February 13, 2003), “All immigrants to have compulsory HIV tests as cases rise”, by Oliver Wright and Anthony Browne.

The NZZ am Sonntag (18 January, 2004), “Alle Asylbewerber sollen zum HIV-Test: Das Bundesamt fuer Gesundheit reagiert auf hohe Zahl infizierter Asylsuchender aus Schwarzafrika”.

Thomas, M., S. Schlegel, and M. Hereward (2001), “UNICEF Speaking out! Voices of children and adolescents in East Asia and the Pacific”, *ed. UNICEF South East Asia Regional Office*, IBSN 974-685-012-1.

UNAIDS/WHO Epidemic update (2002), ISBN 92 9173 304 0, Geneva, Switzerland (www.unaids.org).

UNAIDS (2003) AIDS Epidemic update, December 2003, *ed. UNAIDS and WHO*, Geneva, Switzerland.

UNAIDS Epidemiological Fact Sheets updates (to 2003), *UNAIDS*, Geneva, www.unaids.org

Vanlandinghan, M., S. Suprasert, N. Grandjean and W. Sittitrai (1995), “Two views of risky sexual practices among northern Thai males: The Health Belief Model and the theory of Reasoned Action”, *Journal of Health and Social Behaviour*”, Vol.36, 195-212.

Velazquez-Campoy, A., S. Vega, and E. Freire, (2002), “Amplicification of the effects of drug resistance mutations by background polymorphism in HIV-1 protease from African Sub-types” (2002), *Biochemistry*, Vol.41(27), 8613-8619.

WHO’s HIV/AIDS Strategic Framework for South East Asia region 2002-2006, www.who.org

Wain-Hobson, S. (1998), “More ado about HIV’s origins”, *Nature Medicine*, Vol.4, 1001-1002.

Zurn, P., P. Taffe, M. Rickenbach, J.P. Danthine (2001), “Social cost of HIV infection in Switzerland: Final report”, *Health Management Institute and DEEP*, University of Lausanne

Appendix A: Tabulation of Data & Macro Aggregates

Table 1: Summary Statistics: Size of Sample by Country and Macro-Economic Indicators (2000,2001)[53].

COUNTRY	Sample Size in Survey (%)	Age in [9,13] category	Age in [14,17] category	HIV/Aids Prevalence in 15-49 age range %	HIV/Aids Prevalence in Female population age 15-24, %	HIV/AIDS Prevalence in Young Male Population in [15-24], %	Rank in Human Development Index [HDI, UNDP, with 1=Norway, 175=S Leon]	GINI Coefficient [0,Equality, 100,Inequality]
AUSTRALIA	500 (4.5)	278	222 (44.4)	0.100	0.002	0.020	4	35.1
CAMBODIA	508 (5.0)	295	213 (41.9)	2.800	2.500	4.700	130	40.4
CHINA	1,900 (18.9)	1,040	860 (45.3)	0.100	0.100	0.100	104	40.3
EAST TIMOR	512 (5.1)	324	188 (36.7)	0.050	0.100	0.100	143	38.0
HONG KONG	300 (3.0)	166	134 (44.7)	0.100	0.050	0.050	26	52.2
MACAO	202 (2.0)	119	83 (41.4)	0.050	0.002	0.040	25	50.0
INDONESIA	1,034 (10.3)	567	467 (45.2)	1.100	1.000	1.000	112	31.7
PDR. LAOS	500 (5.0)	285	215 (43.0)	0.050	0.050	0.050	135	39.0
MALAYSIA	515 (5.1)	254	261 (50.7)	0.400	0.100	0.100	58	49.2
MONGOLIA	500 (5.0)	235	265 (53.0)	0.000	0.000	0.000	117	33.2
MYANMAR	500 (5.0)	238	262 (52.4)	2.200	1.700	2.300	131	48.1
PHILIPPINES	500 (5.0)	302	198 (39.6)	0.100	0.100	0.100	85	46.2
P.N.GUINEA	500 (5.0)	249	251 (50.2)	0.700	0.400	0.330	132	50.9
SINGAPORE	502 (5.0)	265	237 (47.2)	0.200	0.200	0.210	28	48.0
SOUTH KOREA	500 (5.0)	222	278 (55.6)	0.050	0.050	0.010	30	31.6
THAILAND	500 (5.0)	272	228 (45.6)	1.800	1.700	3.110	74	41.4
VIET NAM	600 (6.0)	332	268 (44.7)	0.300	0.300	0.100	109	36.1
TOTAL (%)	10,073	5,443 (54%)	4,630 (46%)	**CH=0.005**	**CH=0.004**	**CH<0.005**	**CH=10**	**CH=26.9**

[53] All Macroeconomic variables are from UN-Based sources, mainly from UNDP Country-Profile web page (www.undp.org), UNAIDS Country Fact Sheet Annual Publications (www.unaids.org), ILO Yearly book of Statistics and Companion (www.ilo.org), and WHO Annual Reports 2000 and 2001. The UNICEF SEA -Children's Survey was conducted during the year 2001.

(continue, Table 1)

COUNTRY	% Living below the National Poverty Line	% With direct access to clean water	% Illiterate Population (as % of pop age >= 15)	% Enrolled in Primary school (base age group of legal primary school age)	% Enrolled in Secondary school (base age group of legal secondary school age)	Median Age in Population	Percentage of Population in age group [0,14]	Percentage of Population in age group [15,19]	% Estimation Child Labour (census based, various years)
AUSTRALIA	0.010	1.00	0.010	1.00	1.00	34	0.21	0.07	0.00
CAMBODIA	0.360	0.30	0.335	0.835	0.149	19	0.42	0.12	0.13
CHINA	0.050	0.75	0.162	0.948	0.700	29	0.25	0.08	0.12
EAST TIMOR	0.410	0.63	0.520	0.750	0.300	18	0.41	0.11	0.45
HONG KONG	0.005	1.00	0.065	1.00	1.00	35	0.18	0.06	0.00
MACAO	0.010	1.00	0.059	0.849	0.623	24	0.23	0.08	0.00
INDONESIA	0.160	0.78	0.127	0.900	0.605	20	0.31	0.10	0.10
PDR. LAOS	0.390	0.37	0.344	0.802	0.267	18	0.43	0.11	0.21
MALAYSIA	0.014	0.90	0.121	0.987	0.699	21	0.35	0.10	0.03
MONGOLIA	0.360	0.60	0.015	0.884	0.549	22	0.34	0.11	0.19
MYANMAR	0.150	0.72	0.150	0.820	0.313	21	0.31	0.09	0.22
PHILIPPINES	0.280	0.86	0.049	0.900	0.509	22	0.37	0.11	0.08
P.N.GUINEA	0.380	0.42	0.630	0.838	0.212	19	0.39	0.10	0.19
SINGAPORE	0.010	1.00	0.075	1.00	1.00	31	0.18	0.06	0.00
SOUTH KOREA	0.010	0.92	0.021	1.00	1.00	30	0.21	0.08	0.00
THAILAND	0.110	0.84	0.043	0.800	0.460	29	0.25	0.09	0.16
VIET NAM	0.340	0.77	0.073	0.967	0.495	20	0.32	0.11	0.09
SWITZERLAND	0.000	1.00	0.000	1.000	0.971	39	0.17	0.06	0.00

(Continue, Table 1).

COUNTRY	Population Life Expectancy at birth	Male Life Expectancy at birth	Female Life Expectancy at birth	Population Healthy life Expectancy at birth	% Loss of healthy life, Males	% Loss of healthy life, Females	Population Growth (2000 to 2001)	Growth of Labour Force Population (2000 to 2001)
AUSTRALIA	79	77	82	71.5	0.09	0.11	0.011	0.012
CAMBODIA	54	53	59	47.1	0.15	0.17	0.020	0.026
CHINA	71	69	73	62.1	0.12	0.13	0.008	0.009
EAST TIMOR	57	59	56	46.1	0.20	0.20	-0.022	0.012
HONG KONG	80	78	84	72.1	0.10	0.11	0.009	0.011
MACAO	79	77	81	71.0	0.10	0.12	0.013	0.016
INDONESIA	66	63	67	57.4	0.11	0.14	0.013	0.022
PDR. LAOS	54	52	56	44.7	0.16	0.19	0.024	0.022
MALAYSIA	73	68	74	61.6	0.13	0.15	0.023	0.032
MONGOLIA	65	61	69	52.4	0.18	0.19	0.010	0.018
MYANMAR	57	56	61	49.1	0.15	0.17	0.013	0.016
PHILIPPINES	70	65	71	59.0	0.12	0.14	0.022	0.023
P.N.GUINEA	57	55	58	46.8	0.15	0.18	0.024	0.025
SINGAPORE	78	75	80	67.8	0.11	0.14	0.021	0.017
SOUTH KOREA	74	71	78	66.0	0.10	0.12	0.008	0.015
THAILAND	69	66	72	59.7	0.13	0.15	0.007	0.011
VIET NAM	69	67	71	58.9	0.13	0.16	0.013	0.017
SWITZERLAND	80	76	83	72.1	0.08	0.11	0.001	Na

(Continue, Table 1).

COUNTRY	Per Capita GPD (in current $ USA, 2001)	Per Capita Aid (in current $ USA, 2001)	Total Health Spending as % of GDP	Total Private Health Spending as % of Total Health Spending	Public Health Spending as % of Total Government Spending	Externally finance health spending as % of Public H Spending	Out Of Pocket (oop) finance as % of total private health spend
AUSTRALIA	19.100	0.0	0.086	0.301	0.168	0.00	0.53
CAMBODIA	280	33.3	0.072	0.912	0.061	0.66	1.00
CHINA	920	1.1	0.045	0.612	0.128	0.06	0.80
EAST TIMOR	520	100.0	0.020	0.900	0.040	0.80	1.00
HONG KONG	24,300	0.5	0.050	0.600	0.040	0.00	0.90
MACAO	14,100	1.2	0.050	0.500	0.050	0.00	0.80
INDONESIA	680	7.2	0.027	0.745	0.033	0.19	0.96
PDR. LAOS	320	45.0	0.041	0.629	0.057	0.12	1.00
MALAYSIA	3,700	1.1	0.025	0.423	0.060	0.02	1.00
MONGOLIA	420	87.6	0.062	0.346	0.147	0.05	0.75
MYANMAR	420	2.6	0.015	0.849	0.039	0.04	1.00
PHILIPPINES	910	7.4	0.036	0.576	0.066	0.07	0.83
P.N.GUINEA	550	38.7	0.039	0.086	0.123	0.22	0.92
SINGAPORE	20,700	0.2	0.036	0.646	0.026	0.00	1.00
SOUTH KOREA	15,100	0.0	0.051	0.538	0.096	0.00	0.77
THAILAND	1,880	4.6	0.039	0.386	0.133	0.01	0.85
VIET NAM	420	18.0	0.052	0.761	0.063	0.05	1.00
Switzerland	38,400	0.0	0.100	0.451	0.104	0.00	0.73

Appendix B: Summary of Knowledge and Risk Behavior Indicators

This Appendix complements Section 3.1 showing selected summary statistics on the category variables 'knowledge' (k) and 'health risk behaviour' (r). Each (k) and (r) are variables with 6 categories ($k = 0,1,2,3,4,5$ and $r = 0,1,2,3,4,5$). For k, a 0 means 'scores zero pieces of knowledge out of five possible bits of information on HIV/AIDS', whereas a 5 means that he or she has successfully identified the 4 possible ways of virus transmission as well as identifying a 'blood test' as only way to know if a person is HIV-positive. Any value between 0 and 5 is a possible combination of correct answers where some alternatives are missing. For r, a 0 means the child is not in contact with any kind of drug-abuse or substance use situation, whereas a 5 implies the child is in direct contact with situations where children smoke, sniff glue, drink alcohol, use illicit drugs and abuse on prescription drugs. Anything in between are combinations of these 5 possibilities but not all of them at once.

Table B1: Percentage Distribution of those correctly identifying a particular knowledge item of HIV/AIDS, and Percentage Distribution of children among contact with High Health Risk situations, for each independent situation. (s.e)

KNOWLEDGE INDICATORS	All Sample (n=4,630)	Western E. (n=1,215)	Low Income E. (n=1,332)	Communist E. (n=2,083)
% With zero Knowledge	35.5 (0.7)	29.5 (1.3)	42.1 (1.4)	34.8 (1.0)
% Know 'Can get HIV if unprotected sex'	53.9 (0.7)	55.8 (1.4)	54.3 (1.4)	52.6 (1.1)
% Know 'Can get HIV if share needles-drug users'	38.7 (0.7)	39.3 (1.4)	33.7 (1.3)	41.7 (1.1)
% Know 'Can get HIV from mother to infant'	16.7 (0.5)	12.1 (0.9)	9.8 (0.8)	23.9 (0.9)
% Know 'Can get HIV from any other blood contact'	4.4 (0.3)	5.4 (0.6)	3.2 (0.5)	4.6 (0.5)
% Know ' to know if a person is HIV need blood test'	22.1 (0.6)	28.1 (1.3)	15.7 (1.0)	22.7 (0.9)
RISK BEHAVIOUR INDICATORS	**All Sample (n=4,630)**	**Western E. (n=1,215)**	**Low Income E. (n=1,332)**	**Communist E. (n=2,083)**
% Not exposed to Health Risk Behaviour	33.0 (0.7)	30.9 (1.3)	30.2 (1.3)	36.0 (1.1)
% Exposed to probability of Smoking	61.4 (0.7)	62.6 (1.4)	66.3 (1.3)	57.6 (1.1)
% Exposed to probability of Alcohol drinking	45.3 (0.7)	53.4 (1.4)	38.4 (1.3)	45.0 (1.1)
% Exposed to probability of sniffing glue	5.8 (0.3)	8.1 (0.7)	6.7 (0.7)	3.9 (0.4)
% Exposed to probability of Illicit drug use	10.7 (0.5)	11.5 (0.9)	20.3 (1.1)	4.1 (0.4)
% Exposed to probability of Prescription Drug abuse	5.0 (0.3)	6.7 (0.7)	6.2 (0.7)	3.3 (0.4)

Table B2: **Percentage Distribution of cumulative number of items: Each row defines a Category of the final variables K and R. (s.e)**

KNOWLEDGE CATEGORIES	All Sample (n=4,630)	Western E. (n=1,215)	Low Income E. (n=1,332)	Communist E. (n=2,083)
% Knowledge = 0	35.5 (0.7)	29.5 (1.3)	42.1 (1.4)	34.8 (1.0)
% Knowledge = 1	21.3 (0.6)	28.6 (1.3)	20.5 (1.1)	17.5 (0.8)
% Knowledge = 2	23.6 (0.6)	23.0 (1.2)	23.3 (1.2)	24.2 (0.9)
% Knowledge = 3	12.6 (0.5)	11.8 (0.9)	8.3 (0.8)	15.8 (0.8)
% Knowledge = 4	5.4 (0.3)	4.9 (0.6)	4.3 (0.6)	6.4 (0.5)
% Knowledge = 5	1.6 (0.2)	2.1 (0.4)	1.5 (0.3)	1.3 (0.2)
RISK BEHAVIOUR CATEGORIES	**All Sample (n=4,630)**	**Western E. (n=1,215)**	**Low Income E. (n=1,332)**	**Communist E. (n=2,083)**
% Risk = 0	33.0 (0.7)	30.9 (1.3)	30.2 (1.3)	36.0 (1.1)
% Risk = 1	23.6 (0.6)	19.9 (1.1)	28.2 (1.2)	22.9 (0.9)
% Risk = 2	31.1 (0.6)	33.7 (1.4)	22.8 (1.2)	35.0 (1.0)
% Risk = 3	8.0 (0.4)	9.3 (0.8)	12.9 (0.9)	4.2 (0.4)
% Risk = 4	2.9 (0.2)	3.7 (0.5)	4.0 (0.5)	1.7 (0.3)
% Risk = 5	1.4 (0.2)	2.5 (0.4)	1.9 (0.4)	0.4 (1.4)

Table B3: **Correlation coefficients between Category values of the variables Knowledge and Health Risk Behaviour (t-values in brackets)**

Correlation Coefficients	$r=0$	$r=1$	$r=2$	$r=3$	$r=4$	$r=5$
$k=0$	+0.15 (+10.1)	+0.04 (+2.60)	-0.09 (-6.04)	-0.11 (-7.70)	-0.05 (-3.54)	-0.04 (-2.75)
$k=1$	-0.05 (-3.12)	-0.02 (-1.08)	+0.05 (+3.60)	+0.01 (+0.91)	+0.002 (+0.15)	-0.002 (-0.13)
$k=2$	-0.05 (-3.30)	+0.02 (+1.43)	-0.02 (-1.23)	+0.05 (+3.33)	+0.04 (+2.81)	+0.02 (+1.23)
$k=3$	-0.07 (-4.50)	-0.04 (-2.39)	+0.04 (+2.89)	+0.05 (+3.60)	+0.04 (+2.71)	+0.05 (+3.08)
$k=4$	-0.02 (-1.45)	-0.03 (-2.00)	+0.04 (+2.69)	+0.03 (+1.89)	-0.02 (-1.24)	-0.003 (-0.23)
$k=5$	-0.03 (-2.28)	-0.02 (-1.18)	+0.04 (+2.87)	+0.03 (+2.23)	-0.02 (-1.48)	-0.02 (-1.01)

Note: Each value of the knowledge category is tested for correlation against each value of the risk category.

Table B4 **Correlation coefficients between Categories Knowledge and Health Risk Behaviour, with correlation examined for country blocks.(t-values in brackets)**

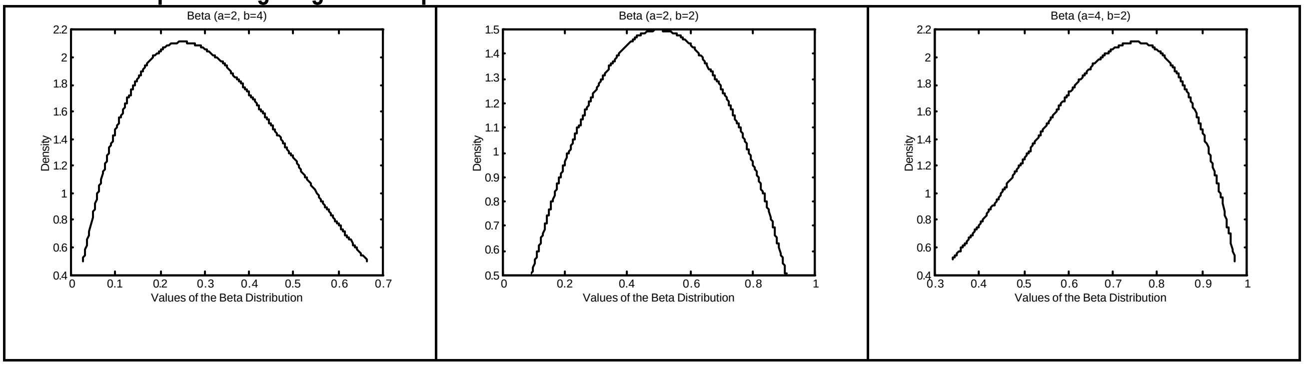
	Correlation Coefficients (knowledge, risk)
All Sample	+0.166 (+11.50)
Western Economies	+0.147 (+5.18)
Low Income Economies	+0.155 (+5.72)
Communist Economies	+0.210 (+9.80)

Note: The variable 'knowledge category' is tested for correlation against the variable 'risk category'.

Appendix C: Workings behind a simulated Probability of inability to avoid HIV-infection (P_A)

Section 3.1 motivates the use of the Beta distribution as an intermediate instrument that classifies each child in the [0,1] interval. The Beta distribution defines a family of probability densities of continuous random variables in the (0,1) interval with $m=(a/a+b)$ and $Var=[ab/(a+b+1)(a+b)^2]$, where a and $b-a$ define location and scale parameters, respectively. If $a=b=1$ the Beta distribution becomes the uniform distribution. Examples of the Beta distribution are given in Table B1.

Table C.1: Representing ranges of shapes for the Beta Distribution.



Each child is classified into one of four possible quadrants (Q.A, Q.B, Q.C, and Q.D) as given in Table 2 (see Section 3.1), according to his or her combination of (k,r), where $k=0,...,5$ and $r=0,...,5$, with low (high) values of k indicating low (high) knowledge, and low (high) values of r indicating low (high) probability of health risk behaviour. Each quadrant is further subdivided into four sub-categories, according to two age groups ([14,15] and [16,17]), and for each of these, combined ascending differences of (k,r). In total, a child can belong to one of 16 possible sub-categories within the two dimensional plane in Table 2 (Section 3.1). The assumption that organizes the16 partitions in ascending order is that knowledge dominates behaviour, while equal (k,r) combination for a younger individuals is valued higher in terms of avoiding HIV-infection.[54] For example, anyone classified in Quadrant A has a (k,r) combination such that $k\geq 3$ and $r\leq 2$. If an individual of age in [14,15] has a combination $(k=5,r=2)$ their P_A value draws

[54] Drawing from Beta does not distinguish between the two age groups if they show perfect knowledge and no risk behaviour, that is, $(k=5,r=0)$.

from a Beta distribution that has a lower mean & smaller variance that an individual in the same age group but with $(k=3, r=2)$. At the same time, if two individual show $(k=3, r=2)$ but each belongs to one of the two age groups, the older individual draws P_A values from a Beta distribution that has slightly higher mean and variance; this assumes that acquiring knowledge and modifying risk behaviour is less likely as children get older. However, within quadrant, the means between selected Beta distributions for the two age groups with equal (k, r) distributions differ only by 0.10 on a [0,1] scale.

Constructing P_A implies taking Beta as defining the relative positions of P_A, given a combination (k, r) and country specific HIV-prevalence ($HIV_{country}$). Based on this, each child's is assigned a Probability of becoming HIV-positive such that $P_A = f(k, r; e) = B(a, b | k, r) + B(a, b | HIV_{country})$. These construction assumes a final outcome variable P_A with values closer to 0 for higher knowledge and low risk behaviour, and values close to 1 for those with very poor knowledge and high probability to engage on risky health behaviour.

Table C2a shows the choice of Beta Parameters (a, b) for each combination in each of the quadrants. Table C2b shows the combination (a, b) that best approximate each country's HIV-prevalence. Random draws of HIV-prevalence adds positive values to each of the 200 draws, for each of the 4,630 children, allowing for country specific additive random shocks to the probability of becoming HIV-positive. In the absence of information on either knowledge or proxy for risk behaviour the HIV-prevalence in each country is the best guide to understand the probability that a young adult becomes infected.[55] Using the Beta specifications in Table C2a, each child draws 200 times from his or her corresponding Beta settings. This implies 200 vectors for P_A, that is, 200 possible set of the outcome of interest. Section 5 and Appendix D show ATE and ATET estimates, for each of the treatment effects, with the final ATE and ATET based on the average over these 200 ATE and 200 ATET values. Here, we are interested on comparing the distribution of P_A for the full sample, between block of countries and between the 17 countries/territories. Thus, to

[55] HIV-prevalence are estimated for the population in the [19,49] age group.

summarize the 200 draws for the sample, the density functions of P_A in Table C3 are based on the per unit average of the 200 draws.

Table C2a: Selection of Beta Distribution, parameters of the Beta Distribution and resulting set of Mean, Variance, by Category (Given Quadrant, Position in Quadrant and Age Group)

		(k,r) combinations	Beta Distributions (a,b)	Quadrant(g, s): (m,s^2)
Quadrant A				
	Best (k,r) **combination**	(5,2), (5,1), (5,0) (4,1), (4,0), (3,0)	**Age [14,15]**: (a=1, b=9) **Age [16,17]:** (a=2, b=12)	**A(1,1)**: (0.10, 0.008) **A(2,1)**: (0.15, 0.008)
	Worst (k,r) **combination**	(4.2), (3,2), (3,1)	**Age [14,15]**: (a=3, b=12) **Age [16,17]**: (a=4, b=12)	**A(1,2)**: (0.20, 0.010) **A(2,2)**: (0.25, 0.011)
Quadrant B				
	Best (k,r) **combination**	(5,5), (5,4), (5,3) (4,4), (4,3)	**Age [14,15]**: (a=5, b=12) **Age [16,17]:** (a=7, b=13)	**B(1,1)**: (0.30, 0.010) **B(2,1)**: (0.35, 0.011)
	Worst (k,r) **combination**	(4,5), (3,5), (3,4), (3,3)	**Age [14,15]**:(a=9, b=13) **Age [16,17]**:(a=10, b=12)	**B(1,2)**: (0.40, 0.010) **B(2,2)**: (0.45, 0.011)
Quadrant C				
	Best (k,r) **combination**	(2,1), (2,0), (1,0)	**Age [14,15]**:(a=12, b=12) **Age [16,17]:**(a=12, b=10)	**C(1,1)**: (0.50, 0.00) **C(2,1)**: (0.55, 0.011)
	Worst (k,r) **combination**	(2,2), (1,2), (1,1) (0,2), (0,1), (0,0)	**Age [14,15]**:(a=12, b=8) **Age [16,17]**:(a=13, b=7)	**C(1,2)**: (0.60, 0.00) **C(2,2)**: (0.65, 0.011)
Quadrant D				
	Best (k,r) **combination**	(2.4), (2,3), (1,3),(0,3)	**Age [14,15]**:(a=14, b=6) **Age [16,17]:**(a=16, b=5)	**D(1,1)**: (0.70, 0.010) **D(2,1)**: (0.75, 0.008)
	Worst (k,r) **combination**	(2,5), (1,5), (1,4) (0,5), (0,4)	**Age [14,15]**:(a=17, b=3) **Age [16,17]**:(a=11, b=1)	**D(1,2)**: (0.85, 0.006) **D(2,2)**: (0.90, 0.006)

Note 1: Knowledge is better risk, so that Quadrant C scores lower Pa (lower chances of becoming HIV positive) relative to where risk is higher but has equal knowledge level.

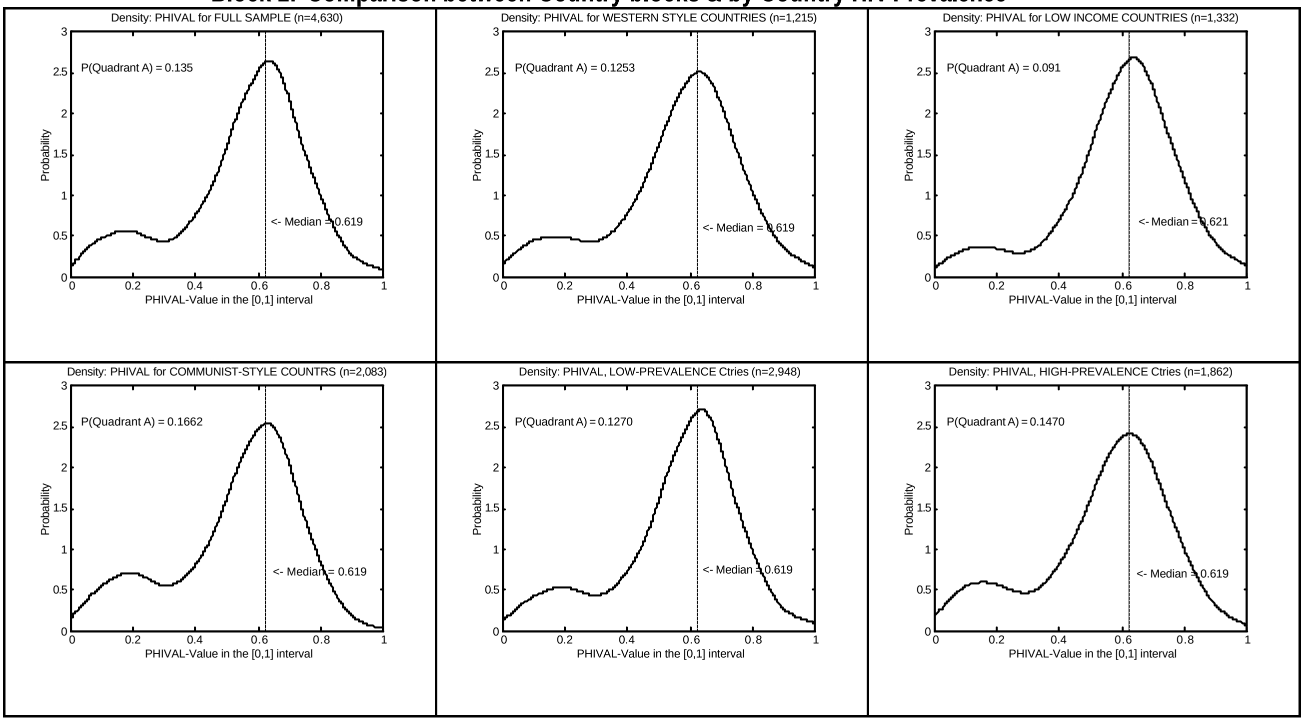
Note 2: The values are selected so that the in each case the variance are not too far apart from 0.10 in the case of the middle values, and slightly smaller in the extreme of the Pa distribution. For example, a Beta distribution with parameters (2,2) also gives a mean of 0.5, but a variance of 0.05. Since we want to set the variance at approximately 0.010, we select the values (12,12) so that the mean is still 0.5 but the variance is 0.01

Table C2b: Selection of Beta Distribution, parameters of the Beta Distribution and resulting set of Mean, Variance, by Category (Given Quadrant, Position in Quadrant and Age Group)

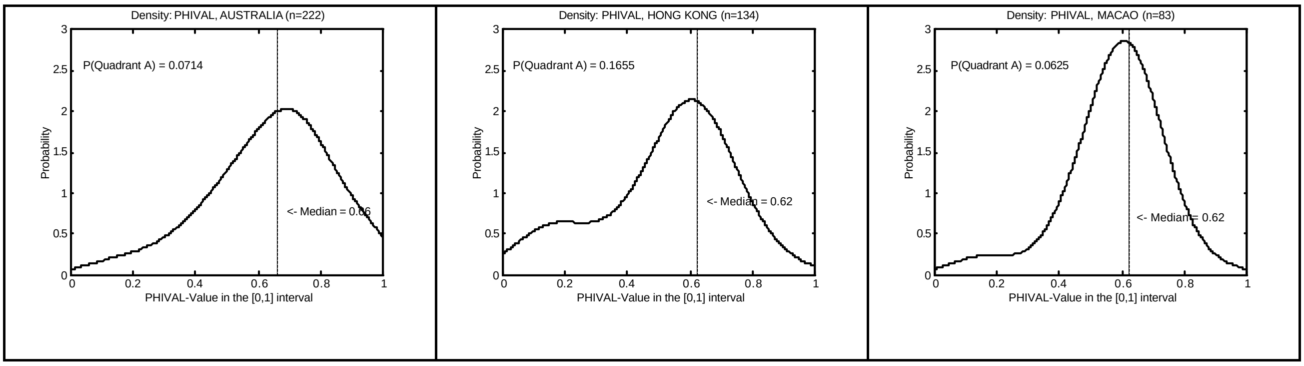
HIV-Prevalence	Countries	Beta Parameters (a,b)
0.0005	Macao, East Timor, South Korea, PDR Laos	B(5,9995)
0.0010	Australia, Hong Kong, Philippines, China	B(1,999)
0.0020	Singapore	B(2,998)
0.0030	Viet Nam	B(3,997)
0.0040	Malaysia	B(4,996)
0.0070	Papua New Guinea	B(7,993)
0.0100	Indonesia	B(1,99)
0.0180	Thailand	B(18,982)
0.0220	Myanmar	B(22,978)
0.0280	Cambodia	B(8,285)
Zero	Mongolia	No Random shocks

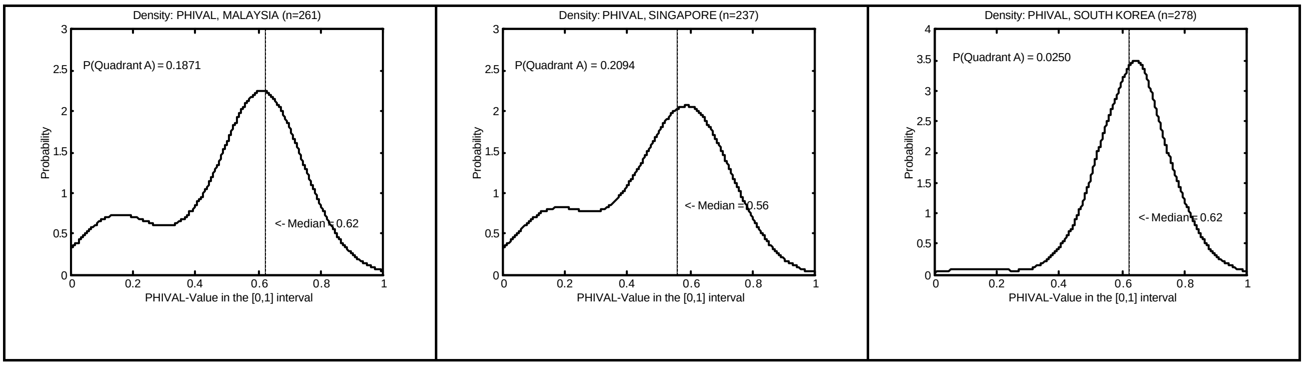
Table C3: Probability Density Functions of P_A

Block 1: Comparison between Country blocks & by Country HIV Prevalence

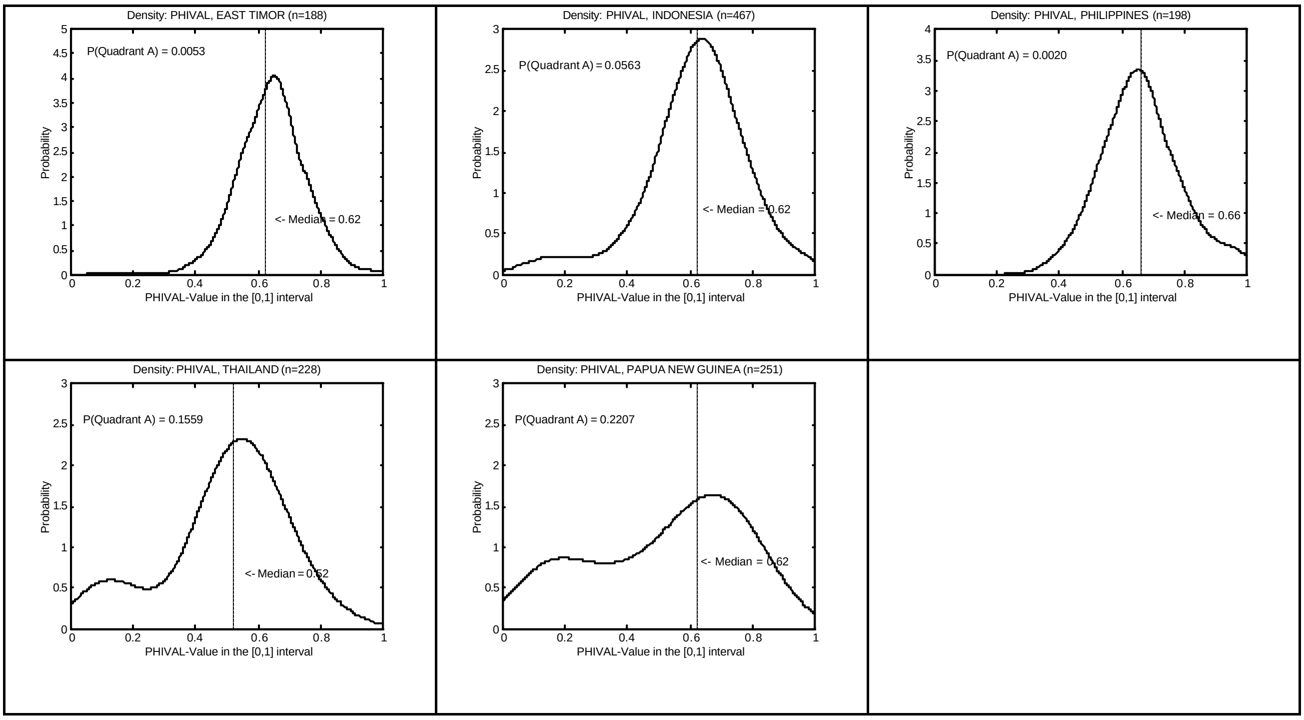


Block 2: Distributions for WESTERN STYLE Countries

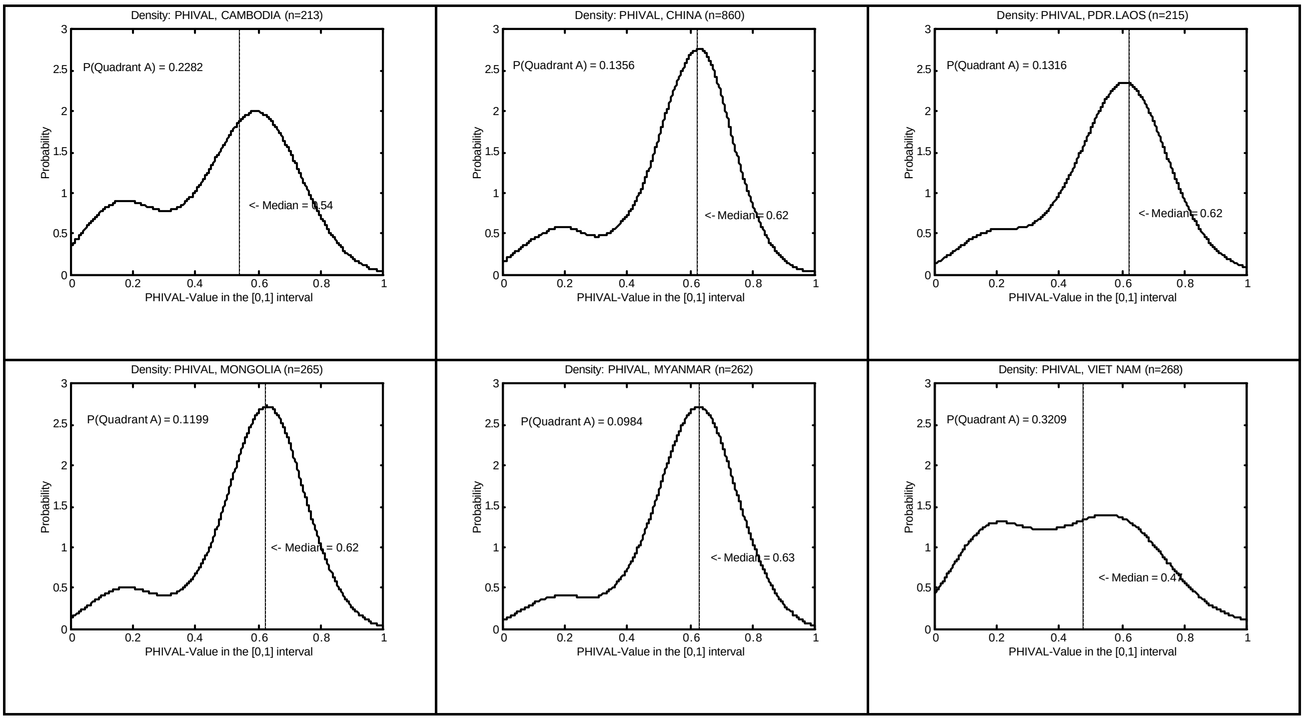




Block 3: Distributions for NON-COMMUNIST LOW INCOME Countries

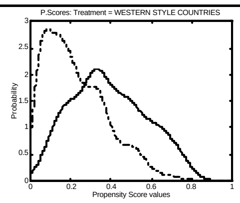


Block 4: Distributions for COMMUNIST STYLE Countries

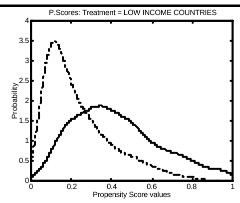


Appendix D: Propensity Score Comparison, ATE and ATET (Estimation based on $P_A = f(k, r; e)$)

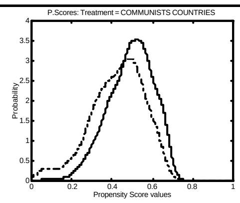
Table D1: Comparison of Treated and Controls Propensity Score density estimates for each of the 43 treatments considered.



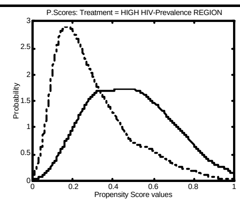
_____ PS treated, ---- PS control



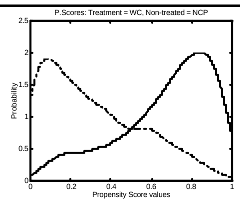
_____ PS treated, ---- PS control



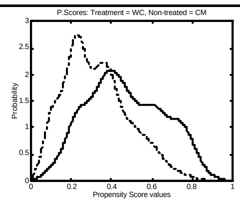
_____ PS treated, ---- PS control



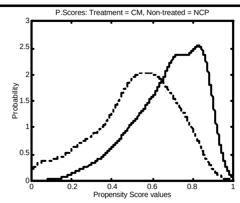
_____ PS treated, ---- PS control



_____ PS treated, ---- PS control



_____ PS treated, ---- PS control



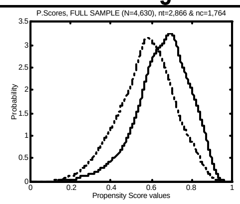
_____ PS treated, ---- PS control

W.C: Western Style Countries.

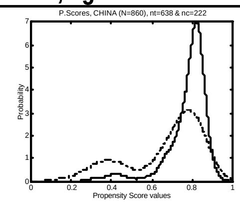
N.C.P: Low Income Countries (i.e., non-communist poor)

CM: Communist style countries

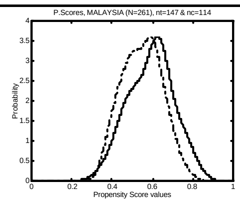
Treatment: Living in Urbanized Area, against Control: Living in Rural Area



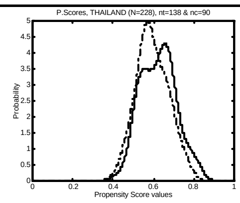
_____ PS treated, ---- PS control



_____ PS treated, ---- PS control

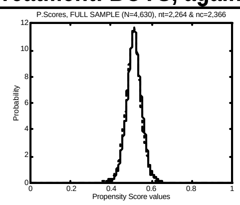


_____ PS treated, ---- PS control

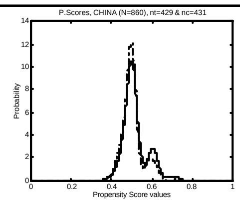


_____ PS treated, ---- PS control

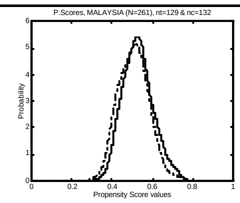
Treatment: BOYS, against Control: GIRLS



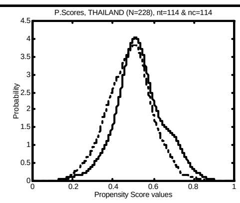
_____ PS treated, ---- PS control



_____ PS treated, ---- PS control

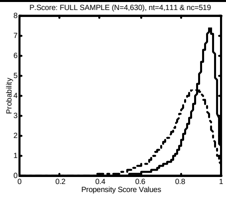


_____ PS treated, ---- PS control

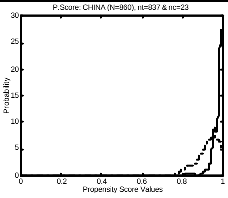


_____ PS treated, ---- PS control

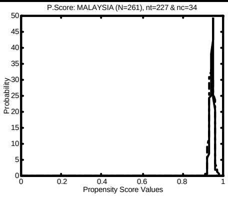
Treatment: AT SCHOOL Status, against Control: NOT AT SCHOOL Status



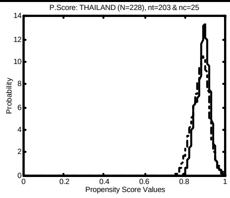
_____ PS treated, ---- PS control



_____ PS treated, ---- PS control

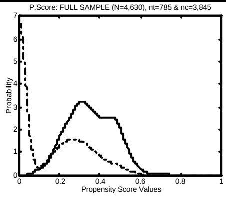


_____ PS treated, ---- PS control

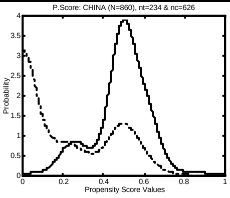


_____ PS treated, ---- PS control

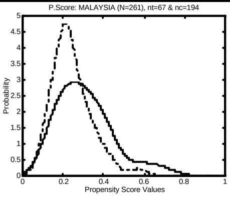
Treatment: HIGH SES Class, against Control: Other two SES Class



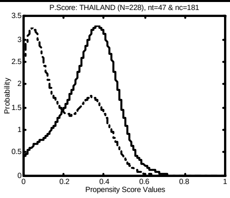
_____ PS treated, ---- PS control



_____ PS treated, ---- PS control

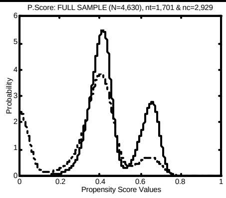


_____ PS treated, ---- PS control

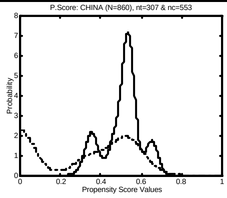


_____ PS treated, ---- PS control

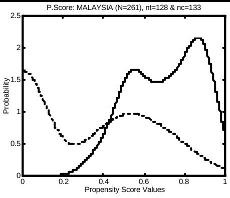
Treatment: MEDIUM SES Class, against Control: Other two SES Class



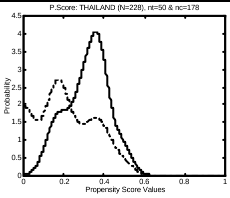
_____ PS treated, ---- PS control



_____ PS treated, ---- PS control

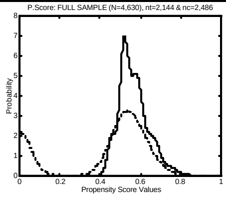


_____ PS treated, ---- PS control

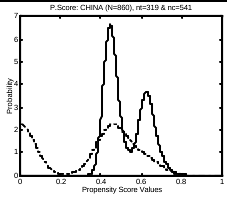


_____ PS treated, ---- PS control

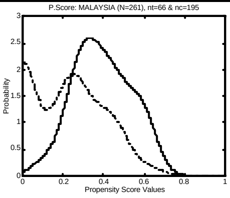
Treatment: LOW SES Class, against Control: Other two SES Class



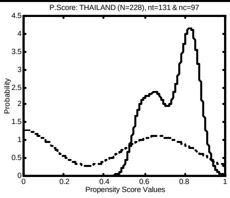
_____ PS treated, ---- PS control



_____ PS treated, ---- PS control

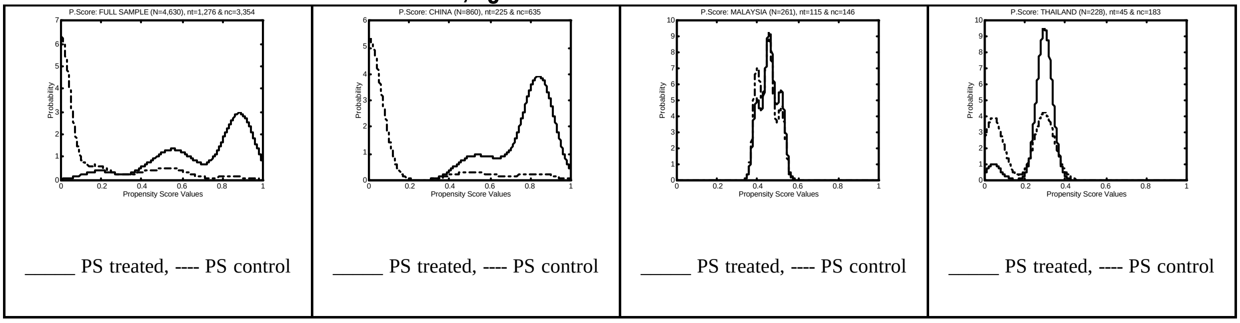


_____ PS treated, ---- PS control

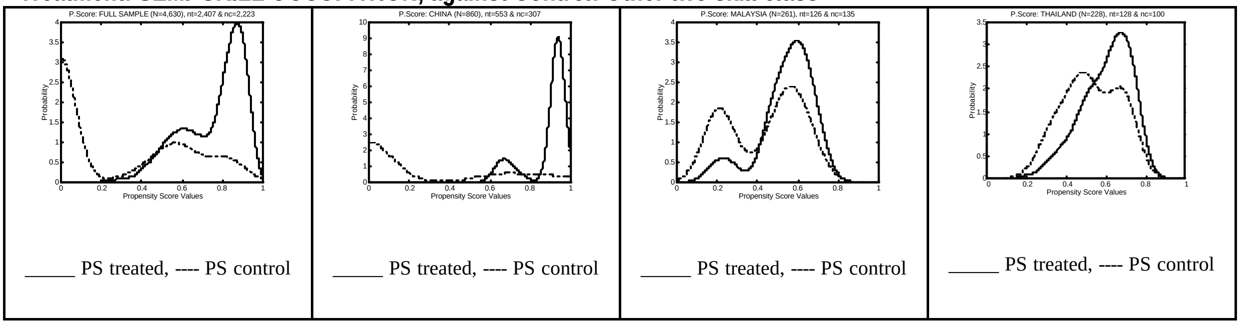


_____ PS treated, ---- PS control

Treatment: PROFESSIONAL OCCUPATION, against Control: Other two skill class



Treatment: SEMI-SKILL OCCUPATION, against Control: Other two skill class



Treatment: UNSKILLED OCCUPATION, against Control: Other two skill class

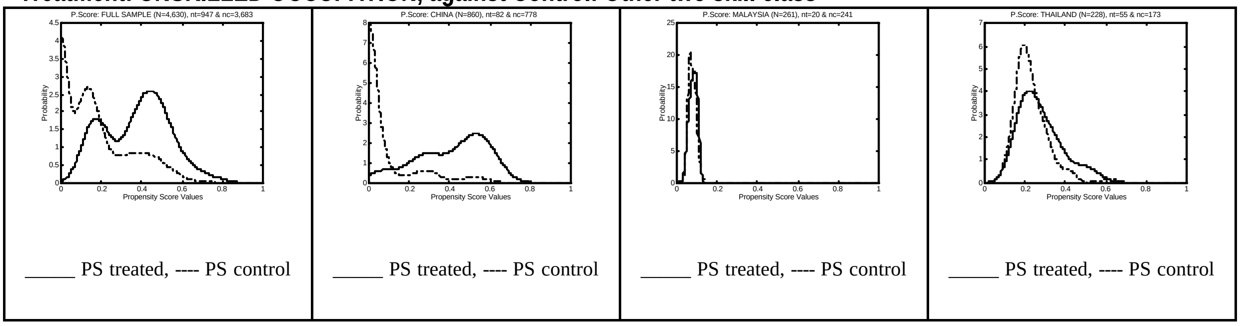


Table D2 and D3 present the results of estimating ATE and ATET, respectively, for each of the 43 independent treatments considered. Each treatment is given a code $(T1,...,T43)$, and in each case the treatment implies a treated and a control sub-sample. Column 2 define these two sub-samples by means of codes. The codes define the following:

ALL:	*Any other unit in the sample of 4,630 that are not included in the treated sample*
WC:	*Western Countries.*
LIC:	*Low Income Countries*
CM:	*Communist Countries*
SES:	*Socio-Economic-Status (SES1: High SES, SES2: Medium SES, SES3: Low SES)*
GTS:	*Go to school Status*
SK:	*Skill level (SK1: Unskilled, SK2: Semi-Skilled, SK3: High Skill or professional)*

Table D2: Average Treatment Effects (ATE), 43 Treatments (Standard errors in brackets), *(0.000) significant at 5%*

Treatment Code	Treated → Control	$E[P_A^T]$ (a)	$E[P_A^C]$ (b)	Δ^T (a)-(b)	95% Confidence Intervals
Geopolitical					
T1	WC → ALL	0.5604 (0.0054)	0.5515 (0.0033)	+0.0089 (0.0063)	(+0.0031, +0.0143)
T2	LIC → ALL	0.5876 (0.0050)	0.5402 (0.0034)	***+0.0473 (0.0060)***	***(+0.0340, +0.0580)***
T3	CM → ALL	0.5285 (0.0043)	0.5746 (0.0037)	***-0.0461 (0.0056)***	***(-0.0593, -0.0317)***
T4	High HIV reg.→ Low	0.5444 (0.0050)	0.5592 (0.0034)	***-0.0148 (0.0060)***	***(-0.0360, +0.0079)***
T5	WC→ LIC	0.5604 (0.0056)	0.5876 (0.0049)	***-0.0271 (0.0073)***	***(-0.0341, -0.0196)***
T6	WC → CM	0.5604 (0.0057)	0.5285 (0.0043)	***+0.0319 (0.0069)***	***(+0.0201, +0.0424)***
T7	CM → LIC	0.5285 (0.0044)	0.5876 (0.0051)	***-0.0591 (0.0068)***	***(+0.0751, +0.0423)***
Socio Economic Treatments					
T8(Full sample)	URBAN → RURAL	0.5463 (0.0036)	0.5661 (0.0044)	***-0.0197 (0.0057)***	***(-0.0350, -0.0040)***
T8(China)	URBAN → RURAL	0.5346 (0.0076)	0.5777 (0.0101)	-0.0431 (0.0255)	(+0.0272, +0.0912)
T8WS(Malaysia)	URBAN → RURAL	0.5360 (0.0167)	0.4758 (0.0204)	***+0.0602 (0.0255)***	***(-0.0264, +0.0136)***
T8(Thailand)	URBAN → RURAL	0.4940 (0.0170)	0.5000 (0.0202)	+0.0063 (0.0265)	(-0.0350, -0.0040)
T9(Full Sample)	Boy → Girl	0.5613 (0.0040)	0.5468 (0.0040)	***+0.0145 (0.0056)***	***(+0.0046, +0.0246)***
T9(China)	Boy → Girl	0.5530 (0.0086)	0.5386 (0.0095)	+0.0143 (0.0131)	(-0.0048, +0.0324)
T9WS(Malaysia)	Boy → Girl	0.5210 (0.0183)	0.4992 (0.0186)	+0.0213 (0.0252)	(-0.0037, +0.0451)
T9(Thailand)	Boy → Girl	0.5045 (0.0187)	0.4881 (0.0180)	+0.0165 (0.0259)	(-0.0017, +0.0371)
T10(Full Sample)	GTS → Not GTS	0.5462 (0.0030)	0.6144 (0.0077)	***-0.0681 (0.0083)***	***(-0.1100, -0.0040)***
T10(China)	GTS → Not GTS	0.5445 (0.0063)	0.5904 (0.0282)	***-0.0460 (0.0290)***	***(-0.1324, +0.0410)***
T10WS(Malaysia)	GTS → Not GTS	0.5071 (0.0142)	0.5567 (0.0480)	-0.0496 (0.0508)	(-0.1062, +0.0096)
T10(Thailand)	GTS → Not GTS	0.4852 (0.0131)	0.5864 (0.0369)	***-0.1012 (0.0323)***	***(-0.1680, -0.0340)***
T11(Full Sample)	SES1 → ALL	0.5530 (0.0074)	0.5541 (0.0032)	-0.0012 (0.0082)	(-0.00082 +0.0067)
T11(China)	SES1 → ALL	0.5431 (0.0127)	0.5467 (0.0071)	-0.0036 (0.0146)	(-0.0184, +0.0082)
T11WS(Malaysia)	SES1 → ALL	0.5020 (0.0273)	0.5124 (0.0154)	-0.0105 (0.0311)	(-0.0315, +0.0116)
T11(Thailand)	SES1 → ALL	0.4575 (0.0302)	0.5064 (0.0143)	-0.0489 (0.0332)	(-0.0752, -0.0189)
T12(Full Sample)	SES2 → ALL	0.5580 (0.0047)	0.5530 (0.0035)	+0.0030 (0.0059)	(-0.0048, +0.0102)
T12(China)	SES2 → ALL	0.5296 (0.0110)	0.5547 (0.0078)	+0.0251 (0.0136)	(-0.0429, -0.0069)
T12WS(Malaysia)	SES2 → ALL	0.5042 (0.0190)	0.5150 (0.0182)	-0.0108 (0.0254)	(-0.0264, +0.0070)
T12(Thailand)	SES2 → ALL	0.5284 (0.0272)	0.4873 (0.0152)	+0.0411 (0.0315)	(+0.0227, +0.0588)
T13(Full Sample)	SES3 → ALL	0.5527 (0.0043)	0.5550 (0.0040)	-0.0021 (0.0060)	(-0.0092, +0.0045)
T13(China)	SES3 → ALL	0.5632 (0.0095)	0.5354 (0.0082)	***+0.0278 (0.0124)***	***(+0.0062, +0.0495)***
T13WS(Malaysia)	SES3 → ALL	0.5283 (0.0227)	0.5034 (0.0162)	+0.0249 (0.0274)	(+0.0019, +0.0459)
T13(Thailand)	SES3 → ALL	0.4980 (0.0168)	0.4940 (0.0205)	+0.0039 (0.0264)	(-0.0196, +0.0265)
T17(Full Sample)	SK1 → ALL	0.5445 (0.0057)	0.5574 (0.0031)	***-0.0130 (0.0066)***	***(-0.0216, -0.0066)***
T17(China)	SK1 → ALL	0.5298 (0.0129)	0.5514 (0.0070)	-0.0216 (0.0144)	(-0.0322, -0.0131)
T17WS(Malaysia)	SK1 → ALL	0.5120 (0.0213)	0.5079 (0.0169)	+0.0042 (0.0266)	(-0.0180, +0.0202)
T17(Thailand)	SK1 → ALL	0.4935 (0.0308)	0.4970 (0.0143)	-0.0035 (0.0337)	(-0.0431, +0.0310)
T18(Full Sample)	SK2 → ALL	0.5545 (0.0038)	0.5532 (0.0040)	+0.0012 (0.0057)	(-0.0050, +0.0073)
T18(China)	SK2 → ALL	0.5473 (0.0076)	0.5429 (0.0104)	+0.0044 (0.0125)	(-0.0082, +0.0166)
T18WS(Malaysia)	SK2 → ALL	0.5048 (0.0182)	0.5143 (0.0189)	-0.0096 (0.0253)	(-0.0260, +0.0094)
T18(Thailand)	SK2 → ALL	0.5141 (0.0161)	0.4735 (0.0210)	+0.0406 (0.0261)	(+0.0142, +0.0680)
T19(Full Sample)	SK3 → ALL	0.5650 (0.0058)	0.5510 (0.0032)	***+0.0140 (0.0068)***	***(+0.0005, +0.0269)***
T19(China)	SK3 → ALL	0.5789 (0.0157)	0.5422 (0.0067)	***+0.0367 (0.0169)***	***(+0.0071, +0.0631)***
T19WS(Malaysia)	SK3 → ALL	0.5275 (0.0399)	0.5082 (0.0143)	+0.0193 (0.0420)	(-0.0115, +0.0452)
T19(Thailand)	SK3 → ALL	0.4571 (0.0283)	0.5087 (0.0144)	-0.0516 (0.0315)	(-0.0722, -0.0334)

Note 1: The expected values $E[P_A^T]$ and $E[P_A^C]$ are based on sample averages of the variable P_A^{group} for each of the treated and control groups, respectively. For sample size in each of the treatments presented see Table 3 in Section 3.2

Note 2: The 95% confidence interval is based on the 5th and 195th values of the 200 estimates of ATE once these are placed in ascending order. The standard error for each measure is based on the average of the 200 standard errors, each in turned based on a bootstrap procedure that draws with replacement from the original sample.

Table D3: Average Treatment Effects on the Treated (ATET), 43 Treatments (Standard errors in brackets) ***(000) significant at 10%, 5.0%***

Treatment Code	Treated → Control	$E[P_A^T \mid T]$ (a)	$E[P_A^C \mid T]$ (b)	Δ^{TT} (a)-(b)	95% Confidence Intervals
Geopolitical					
T1	**WC → ALL**	0.5604 (0.0058)	0.5317 (0.0060)	***+0.0287 (0.0079)***	***(+0.0210, +0.0349)***
T2	**LIC → ALL**	0.5876 (0.0049)	0.5493 (0.0046)	***+0.0383 (0.0066)***	***(+0.0249, +0.0518)***
T3	**CM → ALL**	0.5285 (0.0041)	0.5777 (0.0042)	***-0.0492 (0.0057)***	***(-0.0622, -0.0342)***
T4	**High HIV reg.→ Low**	0.5444 (0.0047)	0.5834 (0.0041)	***-0.0390 (0.0063)***	***(-0.0677, -0.0065)***
T5	**WC→ LIC**	0.5604 (0.0058)	0.5775 (0.0122)	-0.0171 (0.0130)	(-0.0305, -0.0045)
T6	**WC → CM**	0.5604 (0.0056)	0.5183 (0.0071)	***-0.0421 (0.0090)***	***(+0.0294, +0.0524)***
T7	**CM → LIC**	0.5285 (0.0043)	0.5921 (0.0081)	***-0.0637 (0.0093)***	***(-0.0807, -0.0442)***
Socio Economic Treatments					
T8(Full sample)	**URBAN → RURAL**	0.5463 (0.0043)	0.5574 (0.0055)	***-0.0110 (0.0070)***	***(-0.0294, +0.0079)***
T8(China)	**URBAN → RURAL**	0.5346 (0.0076)	0.5821 (0.0106)	***-0.0474 (0.0134)***	***(-0.0771, -0.0149)***
T8WS(Malaysia)	**URBAN → RURAL**	0.5360 (0.0167)	0.4721 (0.0240)	***+0.0639 (0.0276)***	***(+0.0320, +0.0933)***
T8(Thailand)	**URBAN → RURAL**	0.4939 (0.0173)	0.5171 (0.0197)	-0.0232 (0.0268)	(-0.0448, -0.0019)
T9(Full Sample)	**Boy → Girl**	0.5613 (0.0043)	0.5460 (0.0044)	***+0.0153 (0.0061)***	***(+0.0059, +0.0252)***
T9(China)	**Boy → Girl**	0.5529 (0.0080)	0.5429 (0.0100)	+0.0109 (0.0126)	(-0.0131, +0.0332)
T9WS(Malaysia)	**Boy → Girl**	0.5205 (0.0176)	0.4954 (0.0223)	+0.0250 (0.0284)	(-0.0017, +0.0477)
T9(Thailand)	**Boy → Girl**	0.5045 (0.0173)	0.4995 (0.0200)	+0.0050 (0.0266)	(-0.0129, +0.0247)
T10(Full Sample)	**GTS → Not GTS**	0.5462 (0.0029)	0.6166 (0.0098)	***-0.0704 (0.0100)***	***(-0.1097, -0.0316)***
T10(China)	**GTS → Not GTS**	0.5445 (0.0061)	0.5658 (0.0446)	-0.0213 (0.0459)	(-0.1218, +0.0798)
T10WS(Malaysia)	**GTS → Not GTS**	0.5071 (0.0133)	0.5504 (0.0490)	-0.0433 (0.0488)	(-0.0971, +0.0149)
T10(Thailand)	**GTS → Not GTS**	0.4852 (0.0142)	0.5890 (0.0361)	***-0.1038 (0.0404)***	***(-0.1640, -0.0412)***
T11(Full Sample)	**SES1 → ALL**	0.5529 (0.0075)	0.5504 (0.0065)	+0.0025 (0.0098)	(-0.0045, +0.0123)
T11(China)	**SES1 → ALL**	0.5431 (0.0113)	0.5552 (0.0109)	-0.0122 (0.0151)	(-0.0308, +0.0033)
T11WS(Malaysia)	**SES1 → ALL**	0.5019 (0.0267)	0.5226 (0.0189)	-0.0207 (0.0300)	(-0.0426, +0.0042)
T11(Thailand)	**SES1 → ALL**	0.4575 (0.0305)	0.5045 (0.0199)	-0.0470 (0.0368)	(-0.0760, -0.0160)
T12(Full Sample)	**SES2 → ALL**	0.5558 (0.0056)	0.5489 (0.0051)	+0.0069 (0.0080)	(-0.0028, +0.0165)
T12(China)	**SES2 → ALL**	0.5296 (0.0112)	0.5586 (0.0091)	***-0.0290 (0.0146)***	***(-0.0503, -0.0090)***
T12WS(Malaysia)	**SES2 → ALL**	0.5042 (0.0176)	0.5395 (0.0280)	-0.0352 (0.0331)	(-0.6080, -0.0132)
T12(Thailand)	**SES2 → ALL**	0.5284 (0.0254)	0.4963 (0.0207)	+0.0321 (0.0325)	(+0.0017, +0.0531)
T13(Full Sample)	**SES3 → ALL**	0.5527 (0.0043)	0.5563 (0.0057)	-0.0036 (0.0076)	(-0.0122, +0.0044)
T13(China)	**SES3 → ALL**	0.5632 (0.0084)	0.5326 (0.0115)	***+0.0306 (0.0141)***	***(+0.0062, +0.0553)***
T13WS(Malaysia)	**SES3 → ALL**	0.5283 (0.0224)	0.4992 (0.0230)	+0.0291 (0.0296)	(+0.0052, +0.0511)
T13(Thailand)	**SES3 → ALL**	0.4980 (0.0165)	0.5361 (0.0262)	-0.0381 (0.0315)	(-0.0589, -0.0150)
T17(Full Sample)	**SK1 → ALL**	0.5445 (0.0058)	0.5385 (0.0102)	+0.0060 (0.0121)	(-0.0123, +0.0211)
T17(China)	**SK1 → ALL**	0.5298 (0.0126)	0.5671 (0.0224)	***-0.0373 (0.0251)***	***(-0.0555, -0.0182)***
T17WS(Malaysia)	**SK1 → ALL**	0.5120 (0.0217)	0.5024 (0.0184)	+0.0097 (0.0280)	(-0.0077, +0.0238)
T17(Thailand)	**SK1 → ALL**	0.4935 (0.0298)	0.4988 (0.0183)	-0.0053 (0.0337)	(-0.0498, +0.0316)
T18(Full Sample)	**SK2 → ALL**	0.5545 (0.0038)	0.5443 (0.0085)	***+0.0102 (0.0087)***	***(+0.0003, +0.0190)***
T18(China)	**SK2 → ALL**	0.5473 (0.0071)	0.5691 (0.0221)	-0.0218 (0.0241)	(-0.0375, -0.0072)
T18WS(Malaysia)	**SK2 → ALL**	0.5048 (0.0176)	0.5238 (0.0230)	-0.0191 (0.0293)	(-0.0390, +0.0024)
T18(Thailand)	**SK2 → ALL**	0.5141 (0.0153)	0.4776 (0.0238)	+0.0365 (0.0268)	(+0.0072, +0.0660)
T19(Full Sample)	**SK3 → ALL**	0.5650 (0.0058)	0.5617 (0.0050)	+0.0033 (0.0071)	(-0.0090, +0.0164)
T19(China)	**SK3 → ALL**	0.5789 (0.0161)	0.5642 (0.0173)	+0.0147 (0.0234)	(+0.0071, +0.0219)
T19WS(Malaysia)	**SK3 → ALL**	0.5275 (0.0394)	0.5109 (0.0194)	+0.0166 (0.0425)	(-0.0096, +0.0390)
T19(Thailand)	**SK3 → ALL**	0.4571 (0.0264)	0.5055 (0.0188)	***-0.0484 (0.0313)***	***(-0.0690, +0.0304)***

Note 1: The expected values $E[P_A^T \mid T]$ is based on the sample average of the treated. The $E[P_A^C \mid T]$ is based on the sample average of the selected counterfactuals. In each case the average is that of the outcome P_A^{group}, with the sample size determined by the sample size of the treated sample (see Section 3.2 for this value).

Note 2: The 95% confidence interval is based on the 5th and 195th values of the 200 estimates of ATE once these are placed in ascending order. The standard error for each measure is based on the average of the 200 standard errors, each in turned based on a bootstrap procedure that draws with replacement from the original sample.

Appendix E: Analysis ATET on P_{A2} (Supplementary analysis where outcome variable of interest is based only on 'k' [$P_{A2} = B(a,b|k) + B(a,b|HIV_{country})$]).

A first round of estimation (see Section 5), suggest that Children in Communist oriented countries do better than any other country block. A possible reason is that 'risk behaviour' drives these results, with children in non-communist countries more exposed to health risk-behaviour. Based on these, a new outcome variable P_{A2} is simulated, where P_{A2} is also interpreted as the Probability of becoming HIV-infected, but in this case such probability is assumed to be only a function of knowledge and HIV-prevalence in the country, therefore, $P_{A2} = B(a,b|k) + B(a,b|HIV_{country})$. Table E1 shows the choice of values for the Beta distribution that help to construct the outcome variable P_{A2} (Table C2 applies to provide random shock estimates as function of HIV-prevalence in each country). Following the set up in Table E1, each unit in the sample of 4,630 is allowed a distribution of possibilities based on 200 draws each.

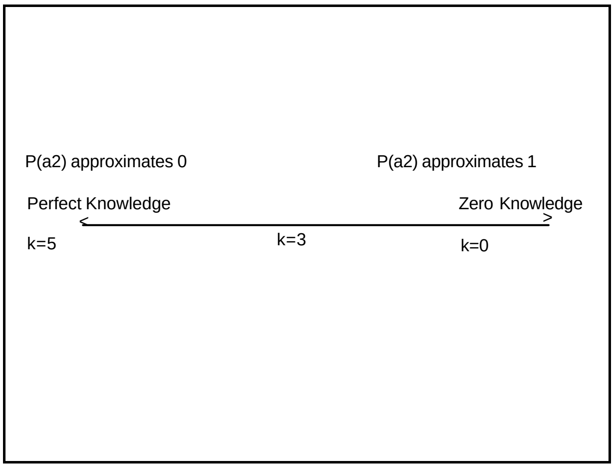
Table E1: Definition of Beta values, mean and variance as a devise to organize sampled children by knowledge values ($k = 0,1,2,3,4,5$)

	(k) Value	Beta Distributions (a,b)	Quadrant(g, s): (m,s^2)
1st Segment in [0,1]		**Age [14,15]**: (a=1, b=19)	**A(1,1)**: (0.05, 0.003)
	5	**Age [16,17]:** (a=1, b=19)	**A(2,1)**: (0.05, 0.003)
2nd Segment in [0,1]		**Age [14,15]**: (a=1, b=9)	**A(1,1)**: (0.10, 0.008)
	4	**Age [16,17]:** (a=2, b=12)	**A(2,1)**: (0.15, 0.008)
3rd Segment in [0,1]		**Age [14,15]**: (a=3, b=12)	**A(1,1)**: (0.20, 0.010)
	3	**Age [16,17]:** (a=4, b=16)	**A(2,1)**: (0.25, 0.015)
4th Segment in [0,1]		**Age [14,15]**:(a=12, b=12)	**A(1,1)**: (0.50, 0.010)
	2	**Age [16,17]:**(a=12, b=10)	**A(2,1)**: (0.55, 0.012)
5th Segment in [0,1]		**Age [14,15]**:(a=14, b=6)	**A(1,1)**: (0.70, 0.010)
	1	**Age [16,17]:** (a=16, b=5)	**A(2,1)**: (0.75, 0.010)
6th Segment in [0,1]		**Age [14,15]**: (a=17, b=3)	**A(1,1)**: (0.85, 0.006)
	0	**Age [16,17]:** (a=10, b=1)	**A(2,1)**: (0.90, 0.008)

Table E2 is analogous to Table 2 (Section 3.1), and interprets the distribution of the values over the [0,1] interval according to $k = 0,...,5$. The % distribution over ascending segments of the [0,1] interval follows the partition in Table E1.

Table E2: **Interpretation of P_{A2} according to k position and P_{A2} percentile distribution over ascending partitions (based on the per unit average over 200 simulated P_{A2} draws).**

Quadrant Distribution of P_{A2}



Distribution of P_{A2} over the [0,1] interval by Beta defined segments.

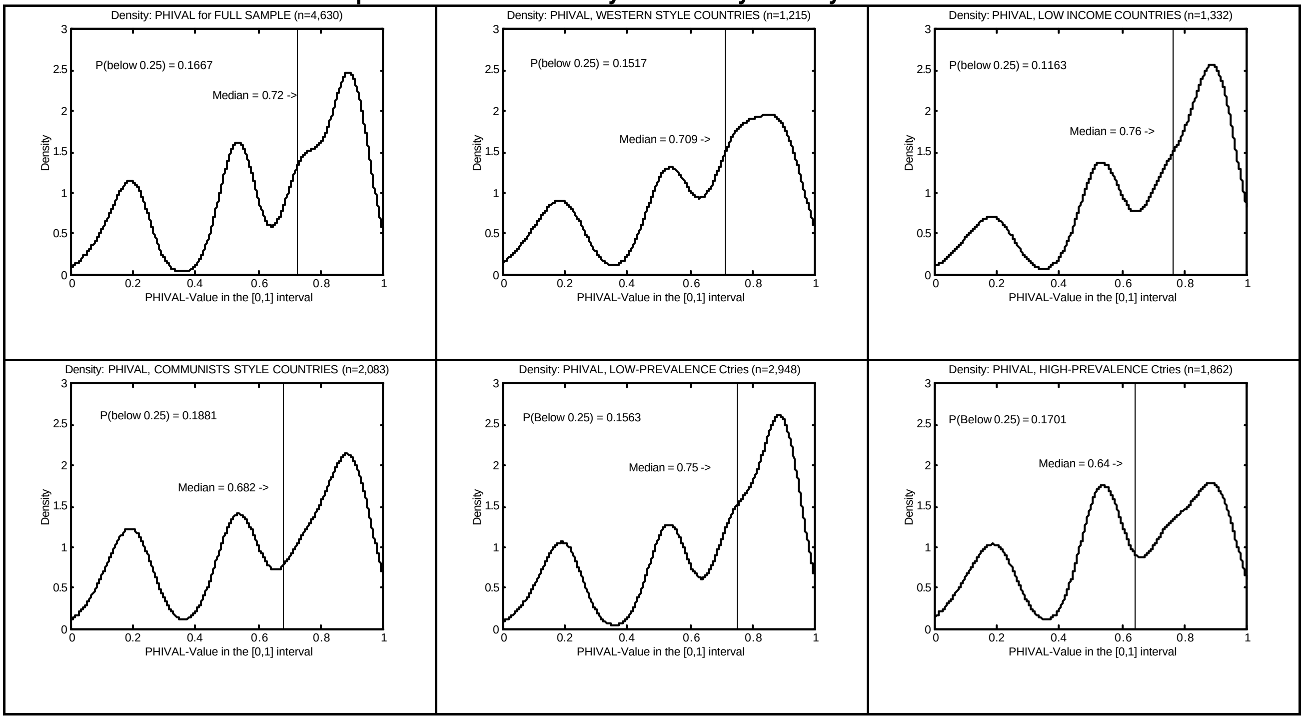
Segments	Percentage (s.e)
% ≤ 0.05	0.09 (0.04)
% $\in (0.05 - 0.15]$	6.63 (0.4)
% $\in (0.15 - 0.25]$	12.96 (0.5)
% $\in (0.25 - 0.55]$	10.04 (0.4)
% $\in (0.55 - 0.75]$	24.7 (0.6)
% > 0.75	45.6 (0.7)

Note1: Estimates are based on 4,630 unit averages over 200 simulated values.

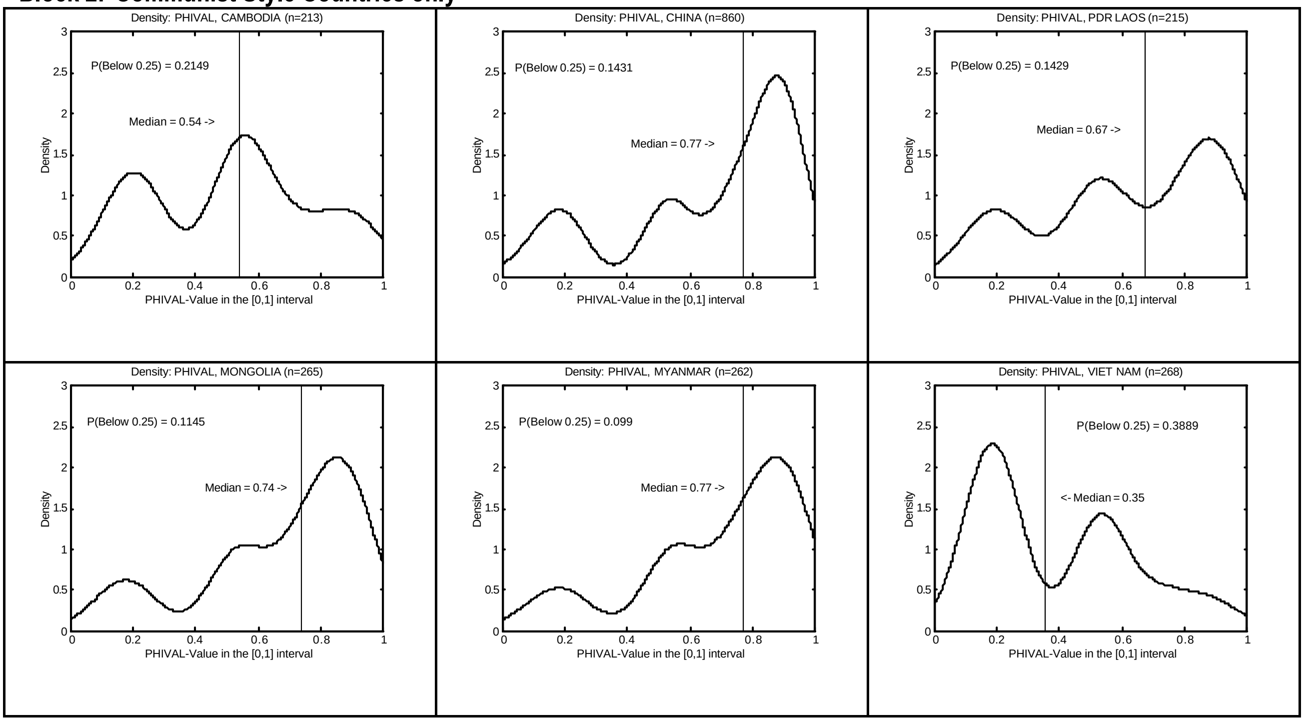
Table E3 shows the smoothed density estimates of P_{A2}. As in Appendix C, this appendix also shows the distribution of P_{A2} for each country-block definition and for each country independently.

Table E3: Probability Density Functions of P_{A2}

Block 1: Comparison between Country blocks & by Country HIV Prevalence



Block 2: Communist Style Countries only



Block 3: Benchmark Comparative Countries: Singapore & Thailand.

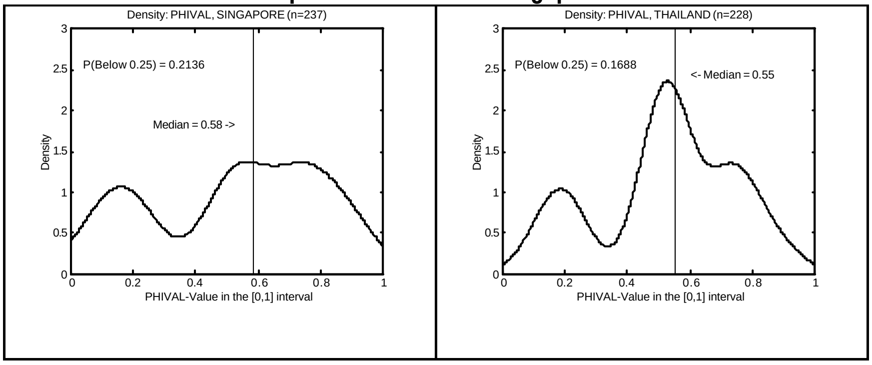


Table E4 shows ATET estimates and intermediate results for a set of treatments where the outcome of interest is P_{A2}. The treatments are defined according to CM (communist block), WC (Western style block), and LIC (Low income countries block).

Table E4: **Average Treatment Effects on the Treated (ATET), 3 Treatments (Standard errors in brackets)** ***(000) significant at 10%, 5.0%***

Treatment Code	Treated → Control	$E[P_{A2}^{T} \mid T]$ (a)	$E[P_{A2}^{C} \mid T]$ (b)	Δ^{TT} (a)-(b)	95% Confidence Intervals
TE1	**CM → ALL**	0.6146 (0.0061)	0.6519 (0.0061)	***-0.0373 (0.0076)***	***(-0.0501, -0.0245)***
TE2	**CM → WC**	0.6146 (0.0061)	0.6522 (0.0089)	***-0.0376 (0.0101)***	***(-0.0492, -0.0230)***
TE3	**CM → LIC**	0.6146 (0.0063)	0.6344 (0.0090)	***-0.0198 (0.0109)***	***(-0.0374, -0.0036)***

Table E5 compares each of the countries in the Communist block separately, to Singapore and Thailand. In this case, the treatment group is always defined by each Communist country independently (CHINA, CAMB (Cambodia), LAOS (PDR Laos), MONG (Mongolia), MYAN (Myanmar), VIET (Viet Nam)). The control groups are always either SING (Singapore) or THAI (Thailand)). In all cases, as in Table E4, the 5th column provides the estimated ATET with corresponding standard errors, and the 6th column are the lower and upper confidence intervals based on the ascending order of the 200 estimated ATET, one for each of the simulated P_{A2} outcomes.

Table E5: **Average Treatment Effects on the Treated (ATET), 3 Treatments (Standard errors in brackets) for each communist country independently against two selected control countries (Singapore and Thailand).*(000) significant at 10%, 5.0%***

Treatment Code	Treated → Control	$E[P_{A2}^{T} \mid T]$ (a)	$E[P_{A2}^{C} \mid T]$ (b)	Δ^{TT} (a)-(b)	95% Confidence Intervals
T-CHINA	**CHINA → SING**	0.6662 (0.0094)	0.5408 (0.0301)	***+0.1254 (0.0315)***	***(+0.0814, +0.1681)***
	CHINA → THAI	0.6662 (0.0097)	0.5447 (0.0260)	***+0.1215 (0.0268)***	***(+0.0550, +0.1799)***
T-CAMBODIA	**CAMB → SING**	0.5170 (0.0200)	0.5740 (0.0342)	***-0.0570 (0.0396)***	***(-0.1068, -0.0044)***
	CAMB → THAI	0.5170 (0.0181)	0.5266 (0.0331)	***-0.0096 (0.0376)***	***(-0.0447, -0.0222)***
T-LAOS	**LAOS → SING**	0.6236 (0.0183)	0.5777 (0.0314)	***+0.0459 (0.0366)***	***(-0.0025, +0.0876***
	LAOS → THAI	0.6236 (0.0198)	0.5306 (0.0253)	***+0.0930 (0.0341)***	***(+0.0496, +0.1343)***
T-MONGOLIA	**MONG → SING**	0.6663 (0.0168)	0.5305 (0.0305)	***+0.1358 (0.0336)***	***(+0.0980, +0.1711***
	MONG → THAI	0.6663 (0.0164)	0.7776 (0.0292)	***-0.1146 (0.0348)***	***(-0.2074, -0.0949)***
T-MYANMAR	**MYNA → SING**	0.6912 (0.0170)	0.5634 (0.0354)	***+0.1278 (0.0391)***	***(+0.0929, +0.1572***
	MYAN → THAI	0.6912 (0.0161)	0.5273 (0.0273)	***+0.1639 (0.0315)***	***(+0.1068, +0.2117)***
T-VIET NAM	**VIET → SING**	0.3935 (0.0154)	0.5643 (0.0247)	***-0.1708 (0.0294)***	***(-0.2163, -0.1249***
	VIET → THAI	0.3935 (0.0143)	0.5308 (0.0297)	***+0.1374 (0.0329)***	***(-0.1779, -0.0930)***