How Economic Development Shapes Household Structure and the Age of Leaving Home and Household Formation: Evidence from 67 countries' 'Big Microdata'

Paavo Monkkonen
Assistant Professor of Urban Planning
UCLA Luskin School of Public Affairs
3250 Public Affairs Building, Box 951656
Los Angeles, CA 90095-1656
E-mail: paavo@luskin.ucla.edu
Telephone: (310) 482-7733

Abstract

The question of whether household structure converges with economic growth and urbanization has been long discussed, but comparative empirical evidence is scant. This paper takes advantage of newly available 'big microdata' to evaluate the relationship between economic growth, urbanization, household structure, the age at which individuals leave their parents' home, and the age at which they form their own household, for almost 70 countries at multiple points in time. To do this, I adapt two measures from demographic research on marriage; the indirect median age and the singulate mean age. I then run cross-section and time series models to examine the correlation to GDP per capita and the urbanization rate. Some findings confirm conventional hypotheses; there is a nearly linear relationship between GDP per capita, the headship rate, share of households that consist of one person, and the share of households made up an extended family. However, the relationship between economic development and the age of leaving home and household formation is not linear; rather, it is a negative quadratic. People leave home and form their own household latest in middle-income countries. Two proposed explanations for this are increasing returns to human capital and changes in the relative cost of housing.

Acknowledgements

1. Introduction

A strong and well-documented connection between economic growth and demographic change, known as the demographic transition, is characterized first by a drop in mortality rates and increase in life expectancy, and then a decrease in fertility rates (Caldwell, 1976). Conventional theories of the past have also equated modernization with a shift in household structure towards smaller, nuclear families, motivating researchers to posit that household structure will change in similar ways as countries develop (Stinner, 1977). Recently, however, this framework has been questioned. Bongaarts (2001), for example, has argued that the transition toward smaller household sizes was erratic in developing countries in the 1990s. Surprisingly, there is limited comparative research that describes the changes in household structure, the age at which individuals form their own household, or the age at which children leave their parental home at a global scale.

Questions about household structure and the age of household formation are important not only for a greater understanding of the cultural changes that occur with economic growth; they are also interconnected to questions of human capital formation, employment, and housing policy. Housing policies are often greatly influenced by calculations of housing needs and housing deficits, which are based on projections of household formation that rely on unquestioned assumptions about household structure and the ages at which individuals leave their parents' home and form their own household (for a discussion, see Monkkonen, 2013). Better empirical examination of this topic can inform public policy, especially housing policy in developing countries.

In this paper, I evaluate the relationship between economic growth and urbanization on the one hand, and household structure and the age at which individuals leave their parents' house on the other. To do this, I use census microdata for over 70 countries at multiple points in time, taking advantage of the 'big microdata' now freely available online (Ruggles, 2013). Thus, the project's scope is larger than other, similar efforts, such as the comparative study of labor and household formation rates in 15 countries (Sevilla-Sanz, 2010). I adapt measures from other areas of demographic research (marriage) to childhood and headship. This is necessary because although census data records individual's household position (child, household head, spouse, relative, grandparent, etc.), they do not record the age at which individuals left their parents' home nor the age at which they formed their own household. To overcome this data challenge, I create two measures of the age at which an "average" person in a country ceased being a child and, separately, formed their own household.

The first of these measures is comparable to a median, and is equivalent to the 'half-life' of childhood. For a given country in a given year, I estimate the age at which 50 percent of people are children. This is a measure similar to the indirect median age of marriage (Riggs, 1953). The second measure is comparable to a mean, and is calculated in the same way as the Singulate Mean Age at Marriage (Hajnal, 1953), which has at least once been applied to the age of leaving home (Steckel, 1996). I also apply these techniques to the age at which an average individual forms their own household, with considerations for gender. Finally (though presented first), I also estimate standard indicators like the headship rate and the share of households that are one-person households and those that include extended family members.

After creating these measures, I examine their relationship to economic development and urbanization, both in a cross-section and over time. This correlation exercise yields important stylized facts, which shed light on debates over the convergence of household structure that is often thought to occur with development. The major finding is an inverted quadratic relationship between the age of leaving home, the age of household formation, and GDP per capita. Several other expected relationships between economic development and household structure are corroborated; such as a strong negative correlation between GDP per capita and the share of households that contain extended family members.

This effort is a first, important step in a research agenda that ought to be extended and with a more detailed spatial focused. As Montgomery (2008) argues, with the continued rapid urbanization of the world's population demographers must increase attention to changes in rapidly urbanizing low- and middle-income countries, especially focusing on small and medium-sized cities. Thus, while this national scale of analysis using 'big microdata' is an important first step, future work should refine the spatial scale of the analysis to assess changes within countries and within cities.

This paper is organized as follows. First, I review the literature on the different factors thought to be associated with the age at which children leave their family home. Then, a description of the data and methods used for analysis is followed by a presentation of the results of the various tests. The conclusion proposes a research agenda around this question, arguing that sub-national data analysis, as well as consideration of migration, is necessary for models to achieve some degree of explanatory power.

2. Literature

The literature on household formation spans a variety of interrelated academic fields, including demography, sociology, economics and theories of urbanization. Four main reasons for leaving the home are identified: entering into a marriage and creating a new family, leaving to obtain education, employment opportunities, and the desire for independence and privacy (De Jong Gierveld et al. 1991). The decision to leave is influenced by numerous socio-demographic factors, with the most prominent influences being income, education, sociocultural norms regarding marriage and family, and housing market conditions. Individuals experience these factors as they intersect with one another rather than in isolation, therefore the ability to identify a sole determining factor in household decision-making is complex. Factors, like higher education, can have the opposite impact depending on the context. In the relatively wealthy United States, for example, many young adults leave the home to obtain higher education, whereas in middle-income countries, getting a university degree is associated with a longer period of 'childhood'.

The literature on this subject primarily focuses on countries where data is most available including Europe, especially the Netherlands and the United States. However, there is a growing body of research surrounding Asian and Latin American countries. The broad differences in demographic trends between developed and developing countries are important to note, although there is evidence of their conversion. Developed countries more typically experience below-

replacement fertility, postponement of legal marriage, an increase in divorce, and rise in the proportion of those living in small households (Burch & Matthews 1987). In developing countries, conversely, urbanization rates are often higher. As a country becomes more urbanized and economically developed, individuals generally have a wider range of opportunities than in previous generations.

Yet, the effect of urban amenities is not uniform across countries. In post-1970s Japan, the availability of urban amenities delays departure from the parental home (Suzuki 2001); in the Netherlands, the degree of urbanization is uncorrelated with the timing of leaving home (Mulder & Hooimeijer 2002); and, in the United States larger urban areas are correlated with earlier departure (Mulder and Clark 2000).

i.   Marriage and Family Formation
Attitudes and views toward marriage and the family are likely the most significant factors in shifting household formation trends across the globe. As countries develop and urbanize, trends observed include: significant delays in marriage (Guest & Tan, 1994; Jones, 2010; Carmichael, 1987), higher divorce rates (Heggeness, 2009; Bonvalet & Lelievre, 1997; Axinn, Amie & Colter, 2008; Jones, 2010), higher rates of women residing alone (Kobrin, 1973); and the increased acceptance of non-marital cohabitation (Axinn, et. al, 2008; Lesthaeghe, 2000; Smits & Mulder, 2008). Goldscheider & DaVanzo (1989) find that decisions are influenced by whether one leaves to reside independently, or to reside with a partner. However, in Latin American countries, De Vos (1989) finds that being married was the major reason for no longer living alone, rather than the educational and income determinants seen in other parts of the world.

The definition and role of family is essential to understanding housing needs (Mulder 2010a). Thornton & Fricke (1987) found significant changes in familial relationships in urbanizing countries, and the concept of family can differ across developed and developing countries. Hajnal (1982) finds that a history of servitude for young adults, which severed them from their families at an early age and led to later marriage, influenced European patterns of household formation. In contrast, in India and China, the tradition of joint households and utilizing extended family rather than servants for household work is likely why trends of co-residence emerged and continue to exist today (Hajnal, 1982). Family norms, more than societal factors, influence marriage formation in developed and developing countries (Billari & Liefbroer; 2007; Buttenheim and Nobles, 2009).

In many Eastern countries the tradition of co-residence, or sharing households with one's parents or other relatives, remains prevalent. Although often times these patterns change as outside economic and urbanization forces shift the focus of individual's lives from the family to outside factors, this is not always the case. The common practice of intergenerational co-residence in east Asia is only marginally affected by level of economic development (Zeng, Coale, Choe, Zhiwu & Li, 1994). Partly due to cultural motivations, married children leave home later than in Western developed countries (Zeng et. al, 1994) One might suspect that the co-residence occurs because of parental need. However, in a study of Indonesia, Singapore and Taiwan, evidence demonstrates that it is the interests of the younger generation that primarily motivate the co-residence (Frankenberg, Chan & Ofstedal, 2002). The Zeng study also suggests that the This is

complemented by research in China demonstrating that co-residence choices in the country are often strategic economic decisions, rather than fixed cultural norms (Logan & Bian, 1999). The historical and economic context of household formation across the world has important implications for present day headship decisions.

Pathway of leaving patterns can also change across countries and are influenced by cultural factors. In a comparative study of Britain and Spain, Holdsworth (2000) found that young Spanish adults are more likely to leave later, and leave for partnership than their British counterparts. This comparative study begins to highlight the importance of familial and cultural norms in household formation decisions.

There are important economic considerations in terms of labor that influence the decision to form a household. In a cross-country study, Sevilla-Sanz (2010) finds that the gender division of labor in a country is influential in household formation decisions, where marriage is seen as two partners contracting over their particular household's future division of labor. Sevilla-Sanz finds that more egalitarian women are less likely to form a household, while more egalitarian men are more likely to do so. This was especially prevalent in low-fertility countries such as Italy, Japan and Spain (Sevilla-Sanz, 2010).

In analyzing patterns of co-residence across 15 developing countries, Ruggles & Heggeness (2008) find that no clear trends exist, and that traditional family forms are stable and even increasing, affirming observations from previous studies discussed above. However, they do find that economic development is positively associated with intergenerational families headed by the older generation (Ruggles & Heggeness 2008). Additionally, Hirschman & Nguyen (2002) find that highly educated men in Vietnam tend to co-reside with their parents, despite their modern education and income level. The authors suggest this could be a reflection of strongly felt familial obligations amongst the most well-educated and elite in the society (Hirschman & Nguyen, 2002). These studies lend further insight to the complex processes of household formation decisions influenced not only by economic factors, but historical and cultural norms, as well.

ii.   Education
While increased levels of educational attainment are often associated with higher levels of independence for young adults, the literature is less decisive. Education level in the Netherlands, for example, is the most important determinant of leaving home (Mulder & Hooimeijer 2002), but evidence from Latin America and Thailand demonstrate the opposite (DeVos 1989; Guest and Tan 1994).

iii.   Income
Headship rates are considerably responsive to income, (Borsh-Supan 1986), particularly in younger populations (Attanasio et al. 2010). In Canada, France, Great Britain, and the United States, the non-family household headship rate has been shown to vary directly with real per capita disposable income (Smith, L., Rosen, K., Markandya, A., and Ullmo, P. 1984). Demographic trends, specifically the aging of the Baby Boomer generation, likely explain the influence of real per capita income on headship rates in the United States from the 1960s to the 1980s (Henderschott and Smith 1984). This can likely be explained by demographic shifts in the

United States including the aging of the 'Baby Boomer' generation. Other hypothesized explanations are the increased demand for privacy as incomes grow (Henderschott and Smith, 1984) or convergence theory—the negative correlation between household size and income in developing countries (Lanjouw and Ravallion 1995). Bongaarts (2001), however, finds that the transition toward smaller household sizes was slow across 43 developing countries in the 1990s, suggesting that factors other than a country's stage of development are influential.

Research has shown that the role of parental income, in particular, can either delay or accelerate children's household formation. Ample parental resources are correlated with earlier departure for educational purposes (De Jong Gierveld, Liefbroer and Beekink 1991) and with later departure for marriage or independent living (Avery, Goldscheider and Speare 1992).

iv.   Relative Housing Prices
Housing markets also shape household formation trends because they impact the affordability of residential options and the ability to form new household. In short, more costly housing can deter people from home ownership (Henrietta 1987) and can reduce household size (Skaburssksis (1994). Household formation patterns are highly responsive to housing prices (Borsh-Supan 1986; Clark & Mulder 2000; Mulder & Clark 2000), rental rates (Haurin, Hendershott & Kim (1993), and the size and regional location of the housing market (Clark & Mulder 2000). In addition, Smith (1984) demonstrates that elasticities of demand are higher for younger age categories. As the relative price of regional housing rises, young adults might choose to delay leaving the home and forming a partnership if they are not able to afford a home of their own (Ermisch 1999). Overall, younger nest-leavers, singles, and urbanites are more likely to share housing and to rent (Clark & Mulder 2000). Unsurprisingly, the increased benefits of homeownership for couples and families make those who leave home for co-habitation for likely to own housing (Mulder 2006). Ermisch & DiSalvo (1996) find that tight housing markets not only discourage young adults from leaving the parental home, especially for women, but also discourage the formation of partnerships for both genders.

Demographic shifts, in turn, can also be traced through housing markets. Population growth rates and changes in the number of homebuyers affect housing prices (Levin, Montagnoli and Wright 2009; Levin et. al., 2009; Mankiw & Weil 1989). In France and Britain, socio-demographic changes such as a decline in birth rate, increase in divorce, and more people living alone has led to a decline in average household size and a corresponding demand for smaller housing options (Bonvalet & Lelievre 1998). In the Netherlands, young singles are attracted to the urban stock of smaller less expensive multifamily units (Kruythoff 1994).

Public policies intended to address housing market problems can have consequences for household formation as well. Welfare and taxation policies shape headship formation around the world, with less developed welfare states often leading to more parental dependence (Haurin, Hendershott & Kim 1993; Lesthaeghe 2000). Housing and land policies, including mortgage guarantees and the tax deductibility of mortgage interest, can influence homeownership and rental rates (Holland, Mulder and Wagner 1998; Pugh 2001).

3. Data and Methods

The data used for this research are from the International Integrated Public Use Microdata Series (IPUMS), a project of the Minnesota Population Center at the University of Minnesota. The IPUMS-International census data contain individual- and household-level data from census bureaus in 74 countries, with most variables harmonized across countries and over time.

The first step in the analysis is to create a set of ten measures of household structure and household formation. I then examine the simple bivariate correlation between these measures and the level of economic development across countries during the most recent census year. The statistical validity of these apparent correlations is tested with cross-sectional models to which I also add the rate of urbanization on the right hand side.

Yet it is difficult to base conclusions on cross-sectional models of relationships that are not static, even in a descriptive sense. Therefore, I run fixed effects models using census data for many countries over multiple years to test how changes in GDP and urbanization affect the age at which individuals leave the family home and other household formation measures. These models control for the time-invariant characteristics of countries. Conclusions based on these models are also necessarily limited, because changes in both variables – economic development and the measures of household structure – change very slowly and in most cases I only have data for a few time periods. Nonetheless, they provide for a robustness check on the more simple cross-sections.

It is important to note that not all countries collect and categorize data in the same way. Inconsistencies in definitions across countries inhibit the ability to compare household structure to some extent. I highlight these limitations where relevant and mitigate them to the extent possible. To ensure consistency, various terms merit definition for this analysis:

i.   Household
A household is defined as one or more persons who make common provision for food, shelter, and other essentials (Bongaarts 2001). Many of the censuses used in this analysis add that households generally share a budget and recognize the authority of a single person. It is important that "family" and "household are not synonymous. In the Philippines, Stinner (1977) found that household sizes are larger in urban than rural areas, but the family size remains the same. In about a dozen countries—including Argentina, Bolivia, Brazil, Chile, Ecuador, Guinea, Iran, Kyrgyz Republic, Mali, Thailand and Uruguay —the census designates households in which none of the members are related as "collective households."

ii.   Headship
The head of a household is defined, most commonly, as the person regarded as such by the household. Each country's census may denote leadership in terms of decision-making, age, or earned income, or general authority. But these definitions vary only slightly and point to a common reference person. Only a few countries have explicit descriptions of headship that dramatically affect the designation. While household heads across the vast majority of countries are explicitly described as a man or woman, the default household head is the male in Canada (1971 only), Ecuador (1962), Hungary (1970, 1980, and 1990), and the United States (1970 only). Furthermore, collective households have no designated "head" in Guinea, Ecuador, and Bolivia.

iii.   Childhood

As a rule, censuses assign all people a household position relative to the household head, be it household head, spouse, child, grandchild, other relative, or non-family member. For the purposes of this study, someone is considered to be a child if she/he is classified as such by the census. This measure is similar to the age at which individuals leave their parental home, but more easily estimated as it relies on household position. "Child" generally includes adopted and step-children. In some censuses there is a distinction. In this paper, all types of children—and grandchildren—are included in the measurement of child status.

There are some small gaps in the data, including a lack of the "child" designation for one census year in one country (Fiji, 1966). A more substantial discrepancy among the data studied is that a few censuses do not include a "grandchild" designation.[1] In such cases, the measure of childhood might be skewed because both children and grandchildren are part of the calculation of the Indirect Mean of Childhood (see below). If there is no "grandchild" category, such household members would be labeled as "other relatives" not included in the count of children.

The ten measures of household structure and household formation trends used in the analysis are adaptations of established demographic indicators, and build upon the above definitions. The definition and construction of each is reported in Table 1, and an elaboration of variables and underlying assumptions is offered afterward.

<<Insert Table 1 here>>

The first five metrics are simple proportions that describe household structure:

i. Headship Rate indicates the percentage of household heads in a population.
ii. Headship Rate with Spouse includes spouses as household heads in the calculation.
iii. One-Person Households are the percentage of households that have one-person.
iv. Extended Family Households are the percentage of households with an extended family member, i.e. related to the household head which are not parent(s) and child(ren).

The next three are singulate mean age indicators, a metric that has been used in demographic research on marriage since the 1950s. The singulate mean age is the number of person-years lived in in a particular demographic category relative to that category's proportion of the population. Hajnal (1953) was the first to apply the singulate means technique to estimate the age of marriage from census data. Numerous demographic studies have since used it for that same purpose (Shyrock 1971; Modell et al. 1976; van Poppel and Derosas 2006). The United Nations (UNDESA 2013) has used the singulate mean age for the past several years in its demographic

---

[1] Countries (and years) in this analysis wherein a census does not include a grandchild category are: Austria (1971), Bangladesh, Bolivia (1992 and 2001), Brazil (1970 and 1980), Burkina Faso (1996), Cameroon (1976 and 1987), Canada (1991 and 2001), Chile (1960), Colombia (1964), Costa Rica (1963 and 1973), Ecuador (1962), France (1962, 1968 and 1975), Guinea (1983), Hungary (1970 and 2001), Indonesia (2000), Kenya (1979, 1989, and 1999), Liberia (1974), Malawi, Malaysia (1970), Mali (1987 and 1998), Mexico (1960, 1970 and 1995), Morocco (1982), Pakistan, Panama (1960), Senegal (1988), Slovenia (2002), Spain (2001), Switzerland, Uganda, Uruguay (1963, 1975, and 1985), Venezuela (1971), Vietnam (1999), Zambia (1990).

reports. Similar to Steckel's (1996) study of 19th century America, I use the singulate mean to describe the age of leaving home. The calculations for the three singulate mean indicators are as follows:

v. Singulate Mean Age at Headship (SMAH) indicates the number of person-years lived as household head relative to the population. It is calculated as:

$$SMAH = (\Sigma_0^{60} S_a - D) / (1 - C)$$

In this equation, $\Sigma_0^{60} S_a$ is the sum of person-years lived not as a household head at age a, $\in$ [0,60]; D is the number of person-years lived by the population that is not household head at age 60; and C is the proportion of population that is not household head at age 60.

vi. Singulate Mean Age at Headship with Spouse (SMAHs) indicates the number of person-years lived as household head or a spouse relative to the population. It is calculated the same way as the above metric but the "D" term covers household heads and spouses.

vii. Singulate Mean Age of Leaving Home (SMALH) is the number of person-years lived as a child/dependent relative to the entire population. It is calculated as:

$$SMALH = \Sigma_0^{100} D_a$$

In this equation, $\Sigma_0^{100} D_a$ is the sum of the proportion of person-years lived by the proportion of children (or grandchildren) at age a, $\in$ [0,100].[2]

The final three indicators are indirect medians, which can be thought of as the half-life of different states. Indirect median in this analysis is the age at which 50% of a group is in position of head or child within a household.[3] Margaret Riggs (1953) first used indirect median in demographic research of personality traits during the 1950s, and subsequent research has yet to apply the metric to headship and childhood in the present day. King and Ruggles (1990) explore the indirect median age of marriage in the United States at the turn of the 20th century; Steckel (1996) and Gjerde (1998) investigate the indirect median age of leaving home in migratory populations in the United States around the same time.

The calculations of indirect medians are as follows:

viii. Indirect Median of Childhood is the age at which 50% of a population is considered children/dependents. It is calculated according to the following formula:

---

[2] According to the IPUMS database, the definition of "child" includes: child (biological status unknown), biological child, adopted child, step-child, child/child-in-law/grandchild, child of unmarried partner. Grandchild includes: grandchild, great-grandchild, great-great grandchild.

[3] There is some potential for measurement error if census enumerators assign household position in different ways in different countries, and for this reason part of the research effort requires a review of differences between census protocols and methods.

$$a_2 - \frac{(a_2 - a_1)(0.5 - r_2)}{(r_1 - r_2)}$$

In this equation, r_1 is the corresponding percentage of children/grandchildren at age a, where $r_1 \in (0.5,1]$; a_2 is the minimum age at which the percentage of children/grandchildren at this age group is smaller than 50%; and r_2 is the corresponding percentage of children/grandchildren at age a_2, where $r_2 \in (0,0.5]$.

ix. Indirect Median of Headship is the age is that at which 50% of a population are household heads. It is calculated as:

$$b_1 + \frac{(b_2 - b_1)(0.5 - t_1)}{(t_2 - t_1)}$$

In this equation, b_1 is the maximum age at which the percentage of household head and spouse at this age group is smaller than 50%.[4] The variable t_1 is the corresponding percentage of household head at age b_1, where $t_1 \in (0,0.5]$, and $t_1 \leq t^*$, where t^* is the maximum percentage of household head. Further, b_2 is the minimum age at which the percentage of household head at this age group is larger than 50%, where $b_2 < b^*$. And t_2 is the corresponding percentage of the proportion of household head at age b_2, where $t_2 \in (0.5,t^*]$.

x. Indirect Median of Headship with Spouse is the age is that at which 50% of a population are household heads or a spouse. It is calculated the same way as the metric above.

I use multiple indicators to ensure the best estimates of the actual phenomenon. The singulate mean and indirect median are two commonly used demographic indicators with advantages and disadvantages when applied to cross-sectional and panel data across many countries. Fitch and Ruggles (2000) argue that household indirect medians for IPUMS data are less sensitive to period effects than singulate means. Hacker (2008) disagrees: because median age is derived from interpolation between two ages, precision is reliant on the number of cases at each age.

4. Results and Analysis

The first part of this section presents some simple scatter plots to visualize the relationship between the ten measures of household structure and household formation, and a country's level of economic development. The second section describes the results of cross-sectional regressions that statistically test these observed relationships. An important challenge in this effort is to capture the long-term process that is change in household structure. A significant cross-sectional, cross-national relationship does not necessarily imply the results of change over time, thus the third section describes the results of panel regressions, which do look at change over time. All scatterplots use the most recent data available from each country, depending on the census year,

---

[4] Moreover, $b_1 < b^*$ where b^* is the age at which the percentage of household head and spouse achieve its maximum t*. I introduce this cut off value because the curve of the percentage of household head and spouse with respect to age will take an upside down U shape, in which the curve display an increasing property with respect to age before reaching its supreme and an decreasing property with respect to age after reaching its supreme. By introducing this cut off value, I guarantee the function is strictly increasing when I impute the indirect median of headship.

and the corresponding GDP for that year. Thus, the data range from the early 1990s to 2010. The regressions pool all available data; over 200 country-year combinations.

Figures 1 and 2 show the positive correlation between two measures of headship rates and economic development. Two types of headship rate are used because the traditional measure primarily reflects household size. Richer countries generally have higher headship rates, which is not surprising given that households have fewer people in them for a number of reasons. Poorer countries generally have higher fertility rates, and there is an assumed greater prevalence of living with extended family member. Such an assumption is borne out in Figures 3 and 4, which show that richer countries have more one-person households and fewer extended family households. The latter relationship is quite important; less than 10% of households are extended family households in the richest countries whereas well over 30% are in the poorest.

<<Insert Figure 1 and 2 here>>>

Figures 1 and 2 show the relationship between traditional the headship rate, the headship calculated including spouses, and the country's GDP per capita. The relationship is essentially linear, indicating that there are more household heads per capita in richer countries. This is not surprising given the fact that household size drops as countries get richer, mostly due to fewer children per household but also greater longevity and more elderly households with one person.

<<Insert Figure 3 and 4 here>>

Figure 3 shows the correlation between the share of households that have only one-person and GDP per capita. The relationship is essentially linear, though it flattens out slightly among poorer countries at a very low share, less than 10 percent of households. Conversely, Figure 4 shows the strong negative correlation between the share of households that contain extended family members, which drops sharply from an average of 30 percent among poorer countries and below 10 percent in most European countries and the United States.

<<Insert Figure 5 and 6 here>>>

Figures 5 and 6 show the relationship between the two measures of the age of leaving home – SMALH and the indeirect median – and economic development. Both exhibit a negative quadratic form, where in low-income countries most individuals leave home before the age of 20, in middle-income countries its 25 or above, and in the higher-GDP countries, like several European countries, the United States, and Canada, it is near 20 years old again.

<<Insert Figure 7 and 8 here>>

The age of household formation also seems to have a quadratic relationship to GDP per capita. Figure 7 and 8 show this relationship. The average age of headship is lowest in poorer countries, peaks in middle-income countries and starts to drop in the wealthiest countries.

In order to ascertain the statistical significance of the relationships observed in Figures 1 – 8, I run cross-sectional regressions of these measures on GDP per capita, its squared value, and the urbanization rate of the country at a given point in time. The models pool all country-year observations available in the IPUMS data. Tables 2 and 3 report the results of these models; the former presents relationships to household structure and the latter to the different measures of the age of leaving home and the age of household formation.

<<Insert Tables 2 and 3 here>>

The models presented in Table 2 confirm a strong and statistically significant relationship between GDP and the array of measures of household structure. Economic development alone explains between 40 and more than 70 percent of the variation in household size, headship rates, and the share of households that have one-person or members of the extended family. Higher-GDP countries have much smaller households, higher headship rates, more one-person households, and fewer households with extended family members. In some cases, such for as headship rates and extended family households, the quadratic equations (using the square of GDP) fit the data slightly better. The urbanization rate of a given country is not associated with any of the household structure variables.

The models in Table 3 confirm the observed, negative quadratic relationship between the age of leaving home and household formation and economic development. Goodness-of-fit measures go up substantially when the quadratic form of GDP is added, for both measures of the age of leaving home and the age of household formation.

Unfortunately, cross-sectional regressions do not reveal trends. It is tempting, but incorrect to assume that the observed relationship between economic growth and household structure and the age of household formation across countries also applies within the same country over time. Trends within a country are not necessarily the same as what is observed between countries (Fields 2001). A further limitation is that cross-sectional analyses of census data depend on the assumption that different cohorts act similarly at each particular age, though this may or may not be the case. One cannot assume that a younger population in one country will evolve to act like an older cohort in another.

To get closer to the dynamic relationship between these variables, I run a second set of regressions that assess at household changes over time in each country. Panel regressions allow a comparison of how household characteristics change in relation to changes in GDP, however, the time period available – in most cases a period of 10 years - is relatively short for such a long-term process. Of the 67 countries for which data are available, 11 only have one year of census data. This means that the time-series models are limited to 56 countries. Data are available for two or three time periods for roughly half of these countries and for the remainder, there are several years' census data. Details of the country-years of data are found in the Appendix.

To test for the negative quadratic relationship found for the age of leaving home and household formation in the cross-section models, I create an interaction term that by multiplying the country's initial GDP by the change in GDP. This term then indicate whether countries that

started at lower-incomes, for example, exhibit a different relationship to changes in GDP than countries that start at middle-income.

<<Insert Tables 4 and 5 here>>

Table 4 reports results from panel regressions of changes in the household structure variables regressed on changes in GDP, changes in the urbanization rate and the interaction term of initial GDP multiplied by changes in GDP. Unlike the strong cross-sectional relationship, changes in GDP occurring at 10-year time periods are mostly not associated with changes in household structure. The only significant relationship is to changes in headship rates, which go up as GDP goes up. Increasing urbanization rates are found to be significantly associated with a decreasing share of one-person households.

Table 5 reports the results of panel regressions of changes in the age of leaving home and household formation on changes in GDP and the urbanization rate. I find no significant relationship between changes in GDP and the age of leaving home in these models, although the signs on the approximation of the quadratic relationship are as expected. On the other hand, there is a statistically significant relationship between changes in GDP and the age of household formation. Interestingly, this is only the case when the approximation of the quadratic form is included. Increases in GDP are associated with increases in the average age at which individuals become a household head, but this is mitigated by the initial GDP of the country, meaning that beyond a certain GDP, increases in GDP are associated with decreases in the age at which individuals become a household head.

5. Conclusions and Further Research

A multitude of factors influence the decision to leave home and to form a new household, the size and structure of households in any given place. Apart from cultural norms about marriage and family structure, wages and the costs of housing are primary factors in a young person's decision to leave home. In some contexts, young people are able to rent an affordable dwelling and leave home, whereas in other cases they are more likely to stay at home and save money until they can purchase a house and move out.

Choices about education are also central to this question. The potential returns to human capital increase with economic development (Schultz, 1961) and this in turn increases individual's incentives to stay in school longer. However, the national and local context of housing costs and parental resources will have a great impact on whether individuals are able to leave home and continue their studies or whether they will stay at home longer in order to increase their human capital.

Given the various avenues through which a country's stage of economic development and urbanization restricts or encourages young adults' ability to leave the parental home and form a household of their own, the strong relationships found in this paper are not surprising. As part of the growing, data-intensive comparative demography literature (Campbell, 2012), the results shed light on debates over the convergence of household structure that has been argued to occur with development. Although the measures of household structure do seem to exhibit a strong

tendency towards smaller households with fewer extended family members, the ages at which households are formed and individuals leave home exhibit greater variation and a non-linear relationship to economic development. The negative quadratic between the age of household formation and economic development is especially intriguing.

Given the broad brush with which this paper paints the phenomena it examines, it is clearly a first step towards a larger research agenda. Future work on this topic should delve into more detail by adding other variables into the models, such as education and potential wages, as well as by increasing the precision of the geographic scope of analysis. National contexts matter for these phenomena, but future work must consider the variation in these indicators within countries, and the variation in the factors - economic opportunities, the benefits of education, cultural norms, and housing costs - that shape them. Migration within countries is also an important consideration. Future work can build on Skaburskis's (1994) work onthe tension between increasing income prospects, which attracts in-migration to cities, and the housing prices that increase with growing urban populations. This paper has demonstrated the potential of now freely available 'big microdata', yet there are many further questions about household formation it can be used to answer.

References

Bongaarts, J. 2001. Household size and composition in the developing world in the 1990s. *Population Studies: A Journal of Demography 55*(3): 263-279.

Borsh-Supan, A. 1986. Household Formation, Housing Prices, and Public Policy Impacts. *Journal of Public Economics 30:* 145-164.

Burch, T. and Matthews, B. 1987. Household Formation in Developed Societies. *Population and Development Review* (13)2: 495-511.

Caldwell, J.C. 1976. Toward a restatement of demographic transition theory. *Population and development review* 321-366.

Campbell, C. 2012. Historical Demography. In Zai Ling, *Demography.* Beijing: Renmin University Press, pp. 233-265.

Ermisch, J. and Di Salvo, P. 1996. Economic Determinants of Young People's Household Formation. *Economica 64:* 627-644.

Fields, G. 2001. <u>Distribution and Development, A New Loos at the Developing World</u>. Russel Sage Foundation, New York, and MIT Press, Cambridge, MA..

Fitch, C. A. and S. Ruggles. 200. "Historical trends in marriage formation: The United States, 1850–1990." <u>The Ties That Bind: Per- spectives on Marriage and Cohabitation</u>. Linda J. Waite, Christine Bachrach, Michelle Hindin, Elizabeth Thomson, and Arland Thornton (Eds.). New York: de Gruyter: 59–88.

Gjerde, J. 1998. "The Minds of the Middle West: Ethnocultural Evolution in the Rural Middle West." University of North Carolina Press.

Hacker, J. D. 2008. "Economic, Demographic, and Anthropometric Correlates of First Marriage in the Mid-Nineteenth-Century United States." *Social Science History,* 3, 23: 307-345.

Hajnal, J. 1953. "Age at Marrying and Proportions Marrying." *Population Sudies Vol. 7: 111-136.*

Haurin, D. R., Hendershott, P. H., & Kim, D. 1993. The impact of real rents and wages on household formation. *The Review of Economics and Statistics* 284-293.

Henderschott, P. and Smith, M. 1984. Household Formations. *Working Paper No. 1390. National Bureau of Economic Research.* Cambridge, MA.

King and Ruggles. 1990. American Immigration, Fertility, and Race Suicide at the Turn of the Century Author(s): Miriam King and Steven Ruggles Source: The Journal of Interdisciplinary History, Vol. 20, No. 3 (Winter, 1990), pp. 347-369 Published by: The MIT Press

Modell, J., F. Furstenberg, and T. Hershberg. 1976. "Social Change and Transitions to Adulthood in Historical Perspective." *Journal of Family History* 1 (1): 7-32.

Monkkonen, P. 2013. Housing Deficits as a Frame for Housing Policy: Housing deficits as a frame for housing policy: demographic change, economic crisis and household formation in Indonesia. *International Journal of Housing Policy* 13: 247-267.

Montgomery, Mark R. 2008. The Urban Transformation of the Developing World. *Science* 319: 761-764.

Mulder, C. and Hooimeijer, P. 2002. Leaving Home in the Netherlands: Timing and first housing. *Journal of Housing and the Built Environment* 17: 237-268.

Mulder, C. and Clark, W. 2000. Leaving Home and Leaving the State: Evidence from the United States. *International Journal of Population Geography* 6: 423-437.

Riggs, M. 1953. "An Investigation of the Nature and Generality of Three New Personality Variables." *Journal of Personality*. Vol. 20, Issue 3, p. 322-344.

Ruggles, Steven. 2012. "The Future of Historical Family Demography." *Annual Review of Sociology*, 38: 423-441.

Ruggles, Steven. 2014. " Big Microdata for Population Research." *Demography,* 51: 287-297

Sevilla-Sanz, A. 2010. Household division of labor and cross-country differences in household formation rates. *Journal of Population Economics* 23: 225–249.

Schultz, Theodore W. 1961. Investment in Human Capital. *The American Economic Review*. 51(1): 1-17.

Shyrock, Henry S. 1971. "The Methods and Materials of Demography." US Government Printing Office.

Skaburskis, A. 1994. Determinants of Canadian Headship Rates. *Urban Studies* 31(8): 1377-1389.

Steckel, R. 1996. "The Age at Leaving Home in the United States, 1850-1860," *Social Science History*, 20, 507–532.

Stern, D. 2004. "The Rise and Fall of the Environmental Kuznets Curve." *World Development* 32 (8): 1419-1435.

Stinner, W. 1977. Urbanization and Household Structure in the Philippines. Journal of Marriage and Family. 39(2) 377-385.

United Nations, Department of Economic and Social Affairs, Population Division. 2013. *World Marriage Data 2012* (POP/DB/Marr/Rev2012).

van Poppel, F. W. A. and Renzo Derosas, Eds. 2006. "Religion and the Decline of Fertility in the Western World."

Table 1. Definitions of Measures of Household Structure and Formation

Metric	Acronym	Definition	Formula
Headship Rate	HR	The percentage of household heads in a population.	t_h/t
Headship Rate with Spouse	HRs	The percentage of household heads or spouses in a population.	t_s/t
One-Person Households	OPH	The percentage of one-person households among all households.	h_o/h
Extended Family Households	EFH	The percentage of extended family households among all households	h_e/h
Singulate Mean Age at Headship	SMAH	Number of person-years lived as household head relative to the population	$(\sum_0^{60} S_a - D\,)\,/\,(\,1 - C\,)$
Singulate Mean Age at Headship with Spouse	SMAHs	Number of person-years lived as household head or spouse relative to the population	$(\sum_0^{60} S_a - D\,)\,/\,(\,1 - C\,)$
Singulate Mean Age of Leaving Home	SMALH	Number of person-years lived as a child/dependent relative to the entire population	$\sum_0^{100} C_a$
Indirect Median of Childhood	IMC	Age is that at which 50% of a population are children/dependents	$a_2 - \dfrac{(a_2 - a_1)(0.5 - r_2)}{(r_1 - r_2)}$
Indirect Median of Headship	IMH	Age is that at which 50% of a population are household heads	$b_1 + \dfrac{(b_2 - b_1)(0.5 - t_1)}{(t_2 - t_1)}$
Indirect Median of Headship with Spouse	IMHs	Age is that at which 50% of a population are household heads or spouses	$b_1 + \dfrac{(b_2 - b_1)(0.5 - t_1)}{(t_2 - t_1)}$

Table 2: OLS Regression for Headship Rate and Household Structure

	Log GDP	Log GDP^2	Urban Rate	Constant	R^2	Observations
Household Size	-1.089***			14.563***	0.393	216
	(0.075)			(0.685)		
Household Size	1.028	-0.125		5.792	0.401	216
	(1.376)	(0.080)		(5.831)		
Household Size	0.09	-0.083	0.014	9.919	0.408	216
	(1.481)	(0.084)	(0.008)	(6.318)		
Headship	0.049***			-0.168***	0.615	212
	(0.002)			(0.020)		
Headship	-0.264***	0.019***		1.129***	0.739	212
	(0.025)	(0.001)		(0.106)		
Headship	-0.256***	0.018***	-0.000	1.095***	0.739	212
	(0.028)	(0.002)	(0.000)	(0.118)		
Headship [a]	0.073***			-0.205***	0.589	212
	(0.003)			(0.030)		
Headship [a]	-0.400***	0.028***		1.756***	0.711	212
	(0.046)	(0.003)		(0.190)		
Headship [a]	-0.384***	0.027***	-0.000	1.687***	0.712	212
	(0.055)	(0.003)	(0.000)	(0.231)		
One-person HH	0.044***			-0.256***	0.459	196
	(0.004)			(0.032)		
One-person HH	-0.340	0.023***		1.332***	0.640	196
	(0.036)	(0.002)		(0.151)		
One-person HH	-0.316***	0.022***	-0.000	1.226***	0.644	196
	(0.040	(0.002)	(0.000)	(0.166)		
Ext. Family HH	-0.060***			0.728***	0.429	193
	(0.005)			(0.039)		
Ext. Family HH	0.207***	-0.016***		-0.376	0.473	193
	(0.066)	(0.004)		(0.0285)		
Ext. Family HH	0.172*	-0.014***	0.000	-0.222	0.477	193
	(0.073)	(0.004)	(0.000)	(0.316)		

Notes: Standard errors are given in parentheses under coefficients. * $p < 0.01$, ** $p < 0.05$, *** $p < 0.01$.
[a] Includes spouse

Table 3. OLS Regression for Age of Leaving Home and Household Formation

	Log GDP	Log GDP^2	Urban rate	Constant	R^2	Observations
SMALH	1.194***			13.105***	0.233	211
	(0.200)			(1.706)		
SMALH	11.702***	-0.623***		-30.459***	0.322	211
	(2.040)	(0.121)		(8.468)		
SMALH	10.811***	-0.582***	0.013	-26.537**	0.325	211
	(2.167)	(0.125)	(0.012)	(9.093)		
Indirect Median of Childhood	0.898***			13.901***	0.215	211
	(0.183)			(1.515)		
Indirect Median of Childhood	8.873***	-0.473***		-19.160**	0.299	211
	(1.1770)	(0.108)		(7.185)		
Indirect Median of Childhood	8.490***	-0.455***	0.006	-17.473*	0.300	211
	(1.793)	(1.108)	(0.009)	(7.346)		
SMAH [a]	1.446***			11.246***	0.223	212
	(0.234)			(2.055)		
SMAH [a]	10.625***	-0.544***		-26.811**	0.267	212
	(2.394)	(0.140)		(10.127)		
SMAH [a]	8.190**	-0.434**	0.035	-16.090)	0.284	212
	(2.751)	(0.154)	(0.019)	(11.655)		
Indirect Median of Headship [a]	-0.0131			26.221***	0.004	211
	(0.187)			(1.563)		
Indirect Median of Headship [a]	10.805***	-0.648***		-19.117**	0.160	211
	(1.612)	(0.097)		(6.656)		
Indirect Median of Headship [a]	9.492***	-0.589	0.019	-13.337	0.171	211
	(1.809)	(0.103)	(0.012)	(7.691)		

Notes: Standard errors are given in parentheses under coefficients. * $p < 0.01$, ** $p < 0.05$, *** $p < 0.01$.
[a] Includes spouse

Table 4: Panel Regression Results: Compound Annual Growth Rate (CAGR) of household structure

	CAGR of GDP	CAGR of Urban.Rate	CAGR of GDP * initial GDP	Constant	R^2	Obs.
CAGR of Avg. Household size	0.414 (0.290)			0.010 (0.017)	0.01	150
CAGR of Avg. Household size	0.504 (0.333)	-1.788 (1.641)		0.028 (0.032)	0.01	150
CAGR of Avg. Household size	-8.361 (6.132)	-0.817 (0.974)	1.162 (0.832)	0.006 (0.020)	0.03	150
CAGR of Headship [a]	0.075*** (0.026)			0.004*** (0.001)	0.03	146
CAGR of Headship [a]	0.076*** (0.026)	-0.019 (0.056)		0.004*** (0.001)	0.03	146
CAGR of Headship [a]	0.098 (0.139)	-0.021 (0.052)	-0.003 (0.019)	0.004*** (0.001)	0.03	146
CAGR % One Person Households	0.058 (0.087)			0.009*** (0.003)	0.01	132
CAGR % One Person Households	0.086 (0.080)	-0.401** (0.171)		0.013*** (0.003)	0.04	132
CAGR % One Person Households	-0.38 (0.672)	-0.357** (0.165)	0.061 (0.083)	0.012*** (0.003)	0.04	132
CAGR % Ext. Family Households	-0.044 (0.198)			-0.002 (0.004)	0.01	130
CAGR % Ext. Family Households	-0.064 (0.198)	0.247 (0.314)		-0.004 (0.005)	0.01	130
CAGR % Ext. Family Households	-0.046 (1.068)	0.245 (0.327)	-0.002 (0.144)	-0.004 (0.006)	0.01	130

Notes: Standard errors are given in parentheses under coefficients. * $p < 0.01$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Panel Regression Results: Compound Annual Growth Rate (CAGR) of the age of leaving home and household formation

	CAGR of GDP	CAGR of Urbanization Rate	CAGR of GDP * initial GDP	Constant	R^2	Observations
CAGR of SMALH	0.021 (0.023)			0.003*** (0.001)	0.003	145
CAGR of SMALH	0.026 (0.025)	-.099 (0.088)		0.004*** (0.001)	0.027	145
CAGR of SMALH	0.264 (0.210)	-0.123 (0.092)	-.031 (0.027)	0.005*** (0.001)	0.039	145
CAGR Ind. Median of Childhood	0.002 (0.018)			0.003*** (0.001)	0.000	145
CAGR Ind. Median of Childhood	0.004 (0.019)	-0.047 (0.049)		0.004*** (0.001)	0.010	145
CAGR Ind. Median of Childhood	0.206 (0.167)	-0.068 (0.048)	-0.026 (0.021)	0.004*** (0.001)	0.026	145
CAGR of SMAH	-0.012 (0.034)			0.006*** (0.001)	0.001	146
CAGR of SMAH	-0.012 (0.034)	-0.011 (0.061)		0.006*** (0.001)	0.001	146
CAGR of SMAH	0.535*** (0.181)	-.066 (0.049)	-.072*** (0.023)	0.007*** (0.001)	0.050	146
CAGR of Ind. Median Headship [a]	0.013 (0.017)			0.002*** (0.001)	0.003	145
CAGR of Ind. Median Headship [a]	0.019 (0.017)	-0.115*** (0.029)		0.003*** (0.001)	0.073	145
CAGR of Ind. Median Headship [a]	0.329** (0.148)	-0.146*** (0.040)	-0.041** (0.019)	0.004*** (0.001)	0.120	145

Notes: Standard errors are given in parentheses under coefficients. $^{*}\,p < 0.01$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$.
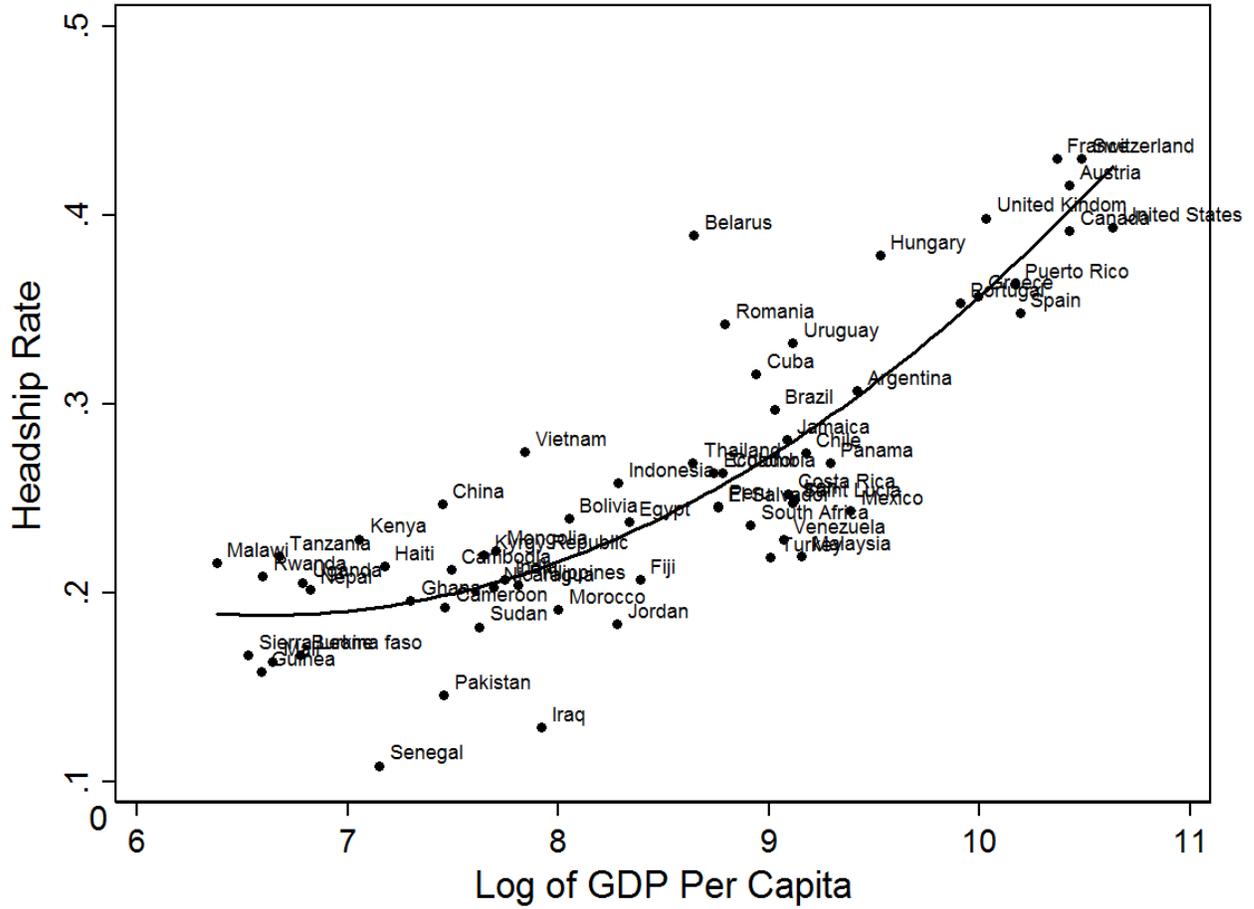
Figure 1: Traditional headship rate and GDP per capita (most recent census)
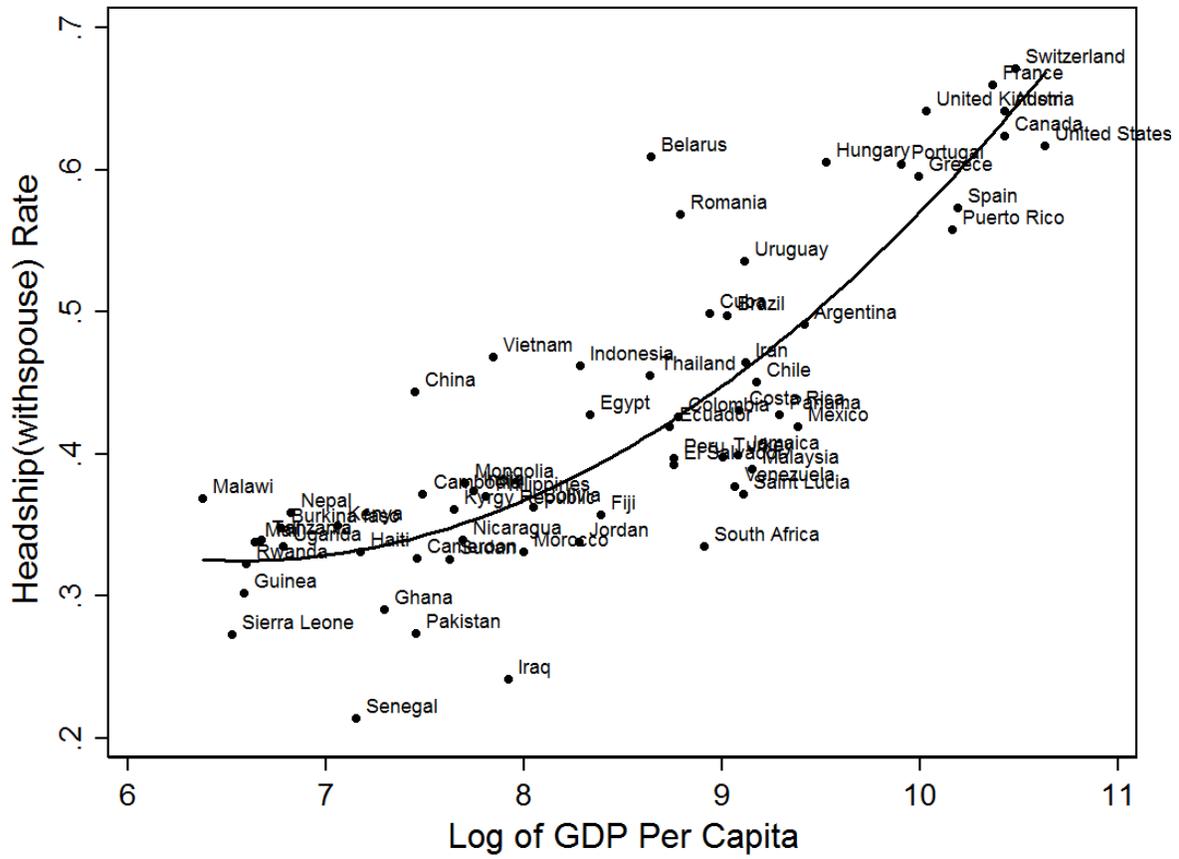
Figure 2: Headship Rate with spouse and GDP per capita (most recent census)
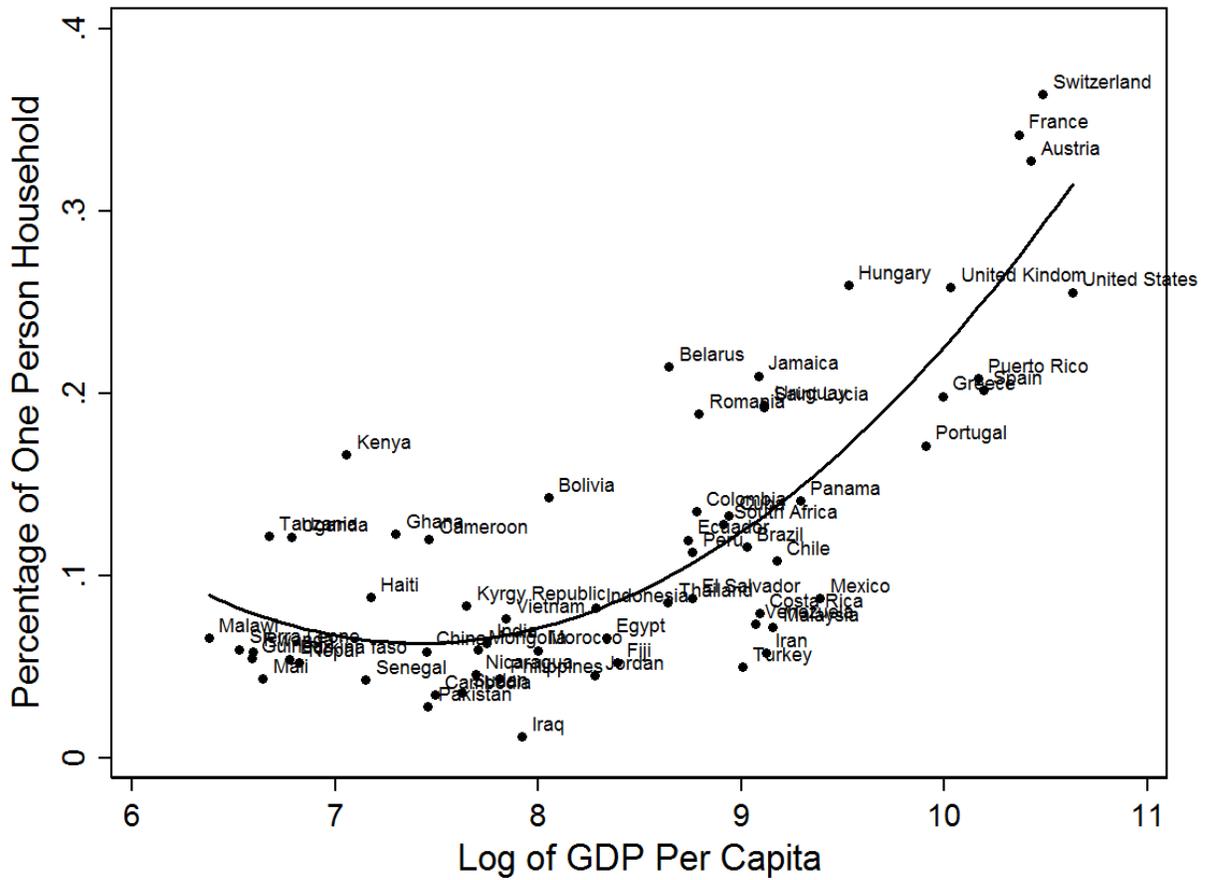
Figure 3: One-person households and GDP per capita (most recent census)
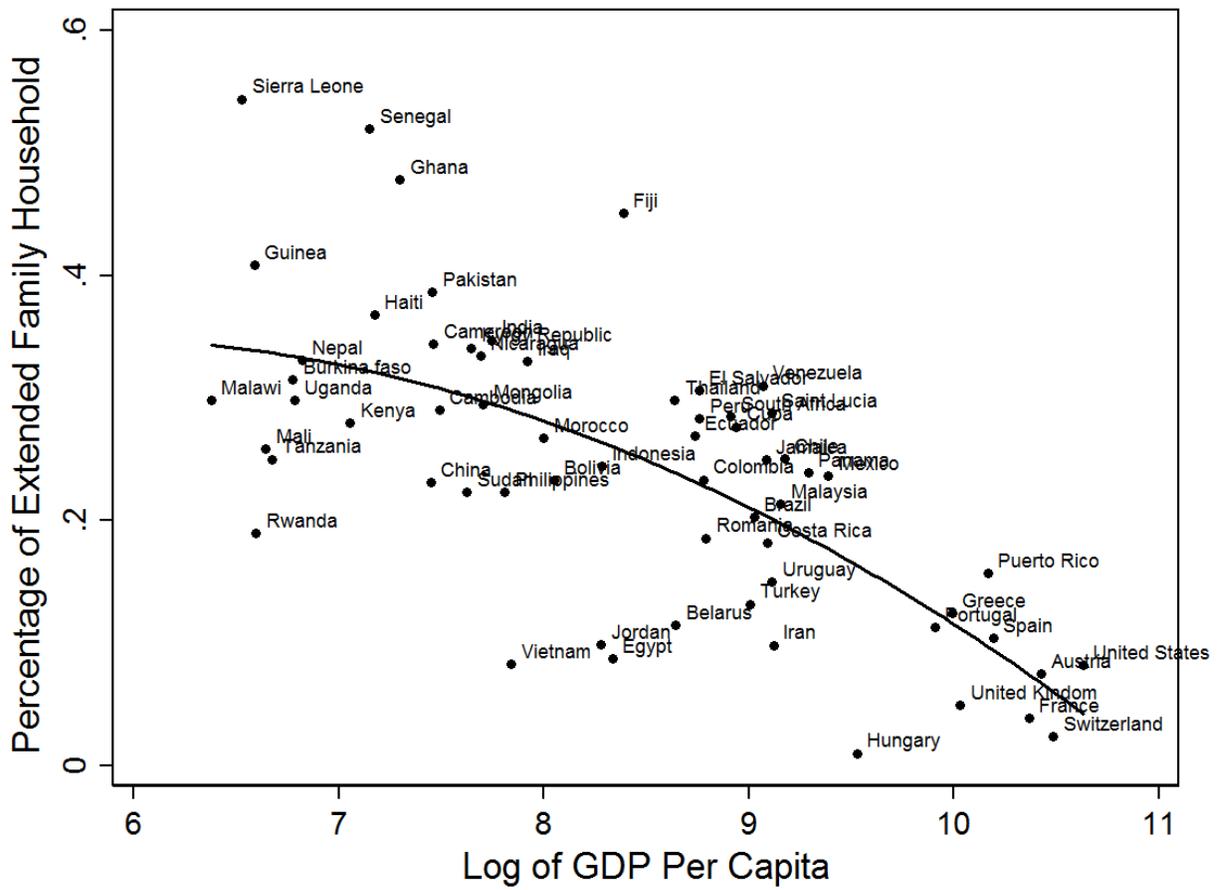
Figure 4: Extended family households and GDP per capita (most recent census)
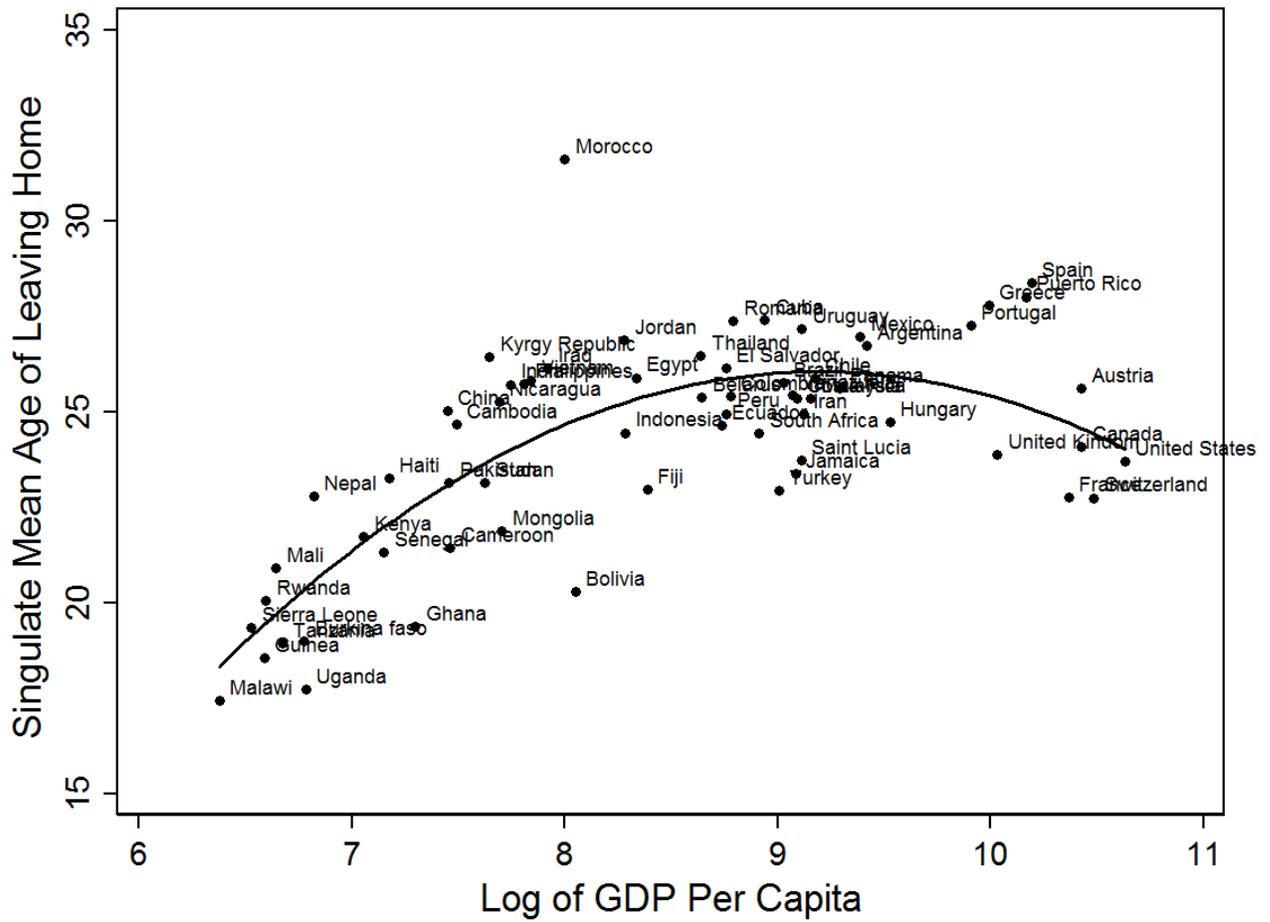
Figure 5: Singulate Mean Age of Leaving Home (SMALH) and GDP per capita (most recent census)
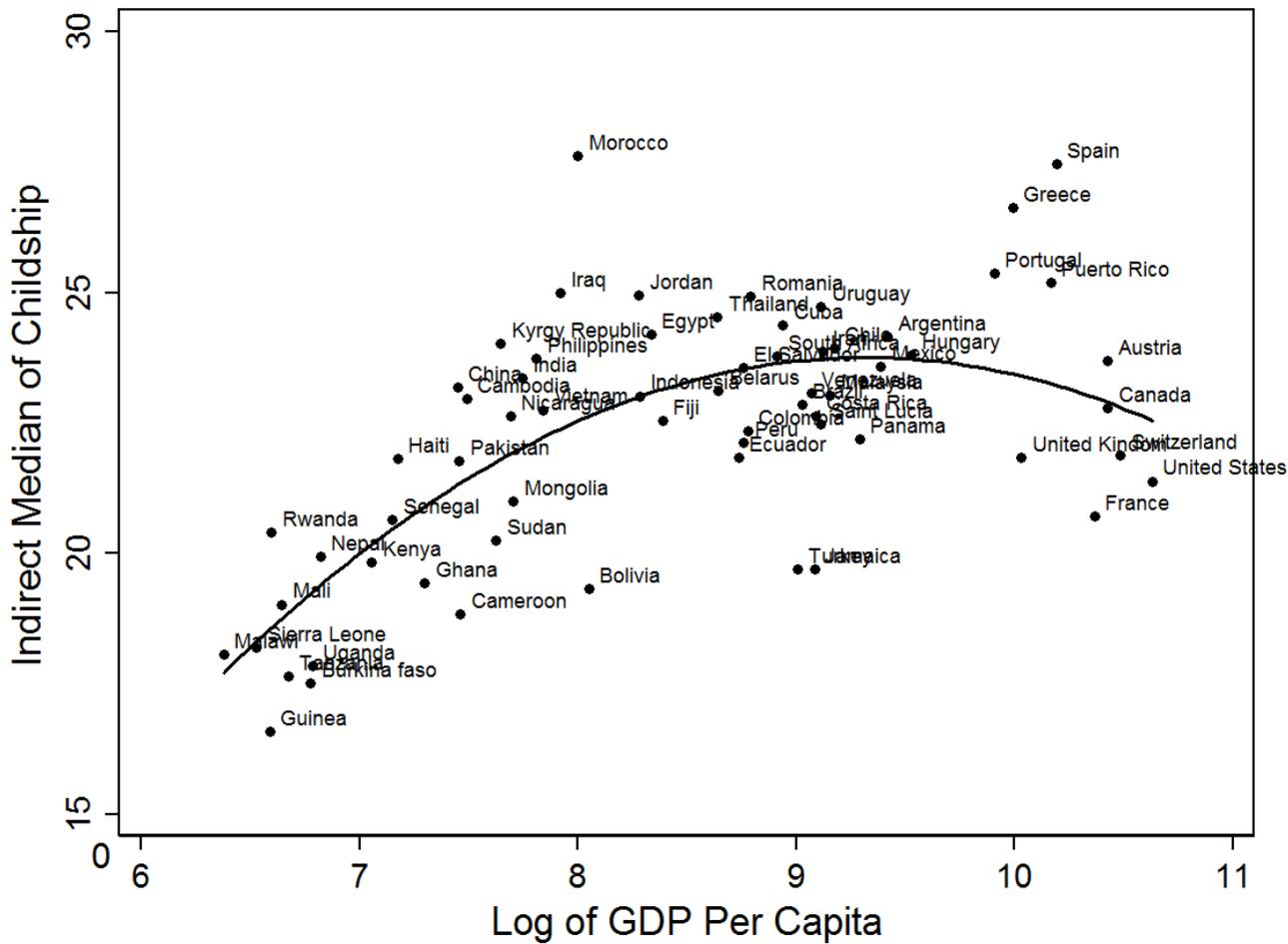
Figure 6: Indirect Median of Childhood and GDP per capita (most recent census)

Figure 7: Singulate Mean Age at Headship (SMAHs) and GDP per capita (most recent census)
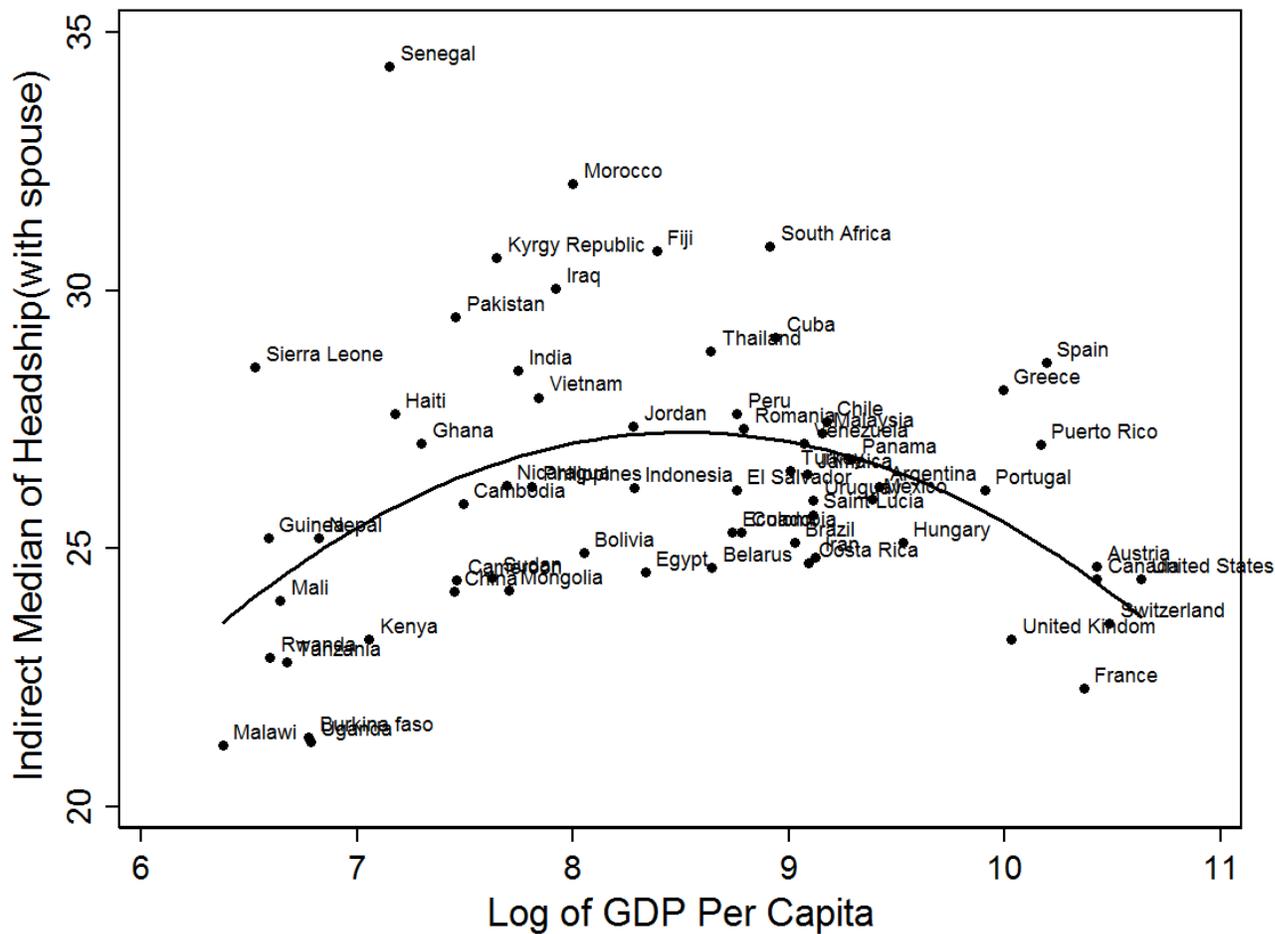
Figure 8: Indirect Median of Headship and GDP per capita (most recent census)

Appendix

In IPUMS International database, there are 74 countries, 238 censuses and 544 million person records in total. In this paper, I exclude Armenia, Ireland, Israel, Italy, Netherlands, Palestine, Slovenia, United Kingdom 2001, because "Age" variable is grouped into categories with different intervals in these samples. Thus, there are 67 countries and 218 censuses in the final analysis. Table 1 shows the countries and years in the sample.

Table 1. Countries and years in the sample

Country	Year
Argentina	1970, 1980, 1991, 2001, 2010
Austria	1971, 1981, 1991, 2001
Bangladesh	1991, 2001, 2011,
Belarus	1999
Bolivia	1976, 1992, 2001
Brazil	1960, 1970, 1980, 1991, 2000, 2010
Burkina Faso	1985, 1996, 2006
Cambodia	1998, 2008
Cameroon	1976, 1987, 2005
Canada	1971, 1981, 1991, 2001
Chile	1960, 1970, 1982, 1992, 2002
China	1982, 1990
Colombia	1964, 1973, 1985, 1993, 2005
Costa Rica	1963, 1973, 1984, 2000
Cuba	2002
Ecuador	1962, 1974, 1982, 1990, 2001, 2010
Egypt	1996, 2006
El Salvador	1992, 2007
Fiji	1966, 1976, 1986, 1996, 2007
France	1962, 1968, 1975, 1982, 1990, 1999, 2006
Germany	1970, 1971, 1981, 1987
Ghana	2000
Greece	1971, 1981, 1991, 2001
Guinea	1983, 1996
Haiti	1971, 1982, 2003
Hungary	1970, 1980, 1990, 2001
India	1983, 1987, 1993, 1999, 2004
Indonesia	1971, 1976, 1980, 1985, 1990, 1995, 2000, 2005, 2010
Iran	2006
Iraq	1997
Jamaica	1982, 1991, 2001

Jordan	2004
Kenya	1969, 1979, 1989, 1999, 2009
Kyrgy Republic	1999, 2009
Malawi	1987, 1998, 2008
Malaysia	1970, 1980, 1991, 2000
Mali	1987, 1998
Mexico	1960, 1970, 1990, 1995, 2000, 2005, 2010
Mongolia	1989, 2000
Morocco	1982, 1994, 2004
Nepal	2001
Nicaragua	1971, 1995, 2005
Pakistan	1973, 1981, 1998
Panama	1960, 1970, 1980, 1990, 2000, 2010
Peru	1993, 2007
Philippines	1990, 1995, 2000
Portugal	1981, 1991, 2001
Puerto Rico	1970, 1980, 1990, 2000, 2005
Romania	1977, 1992, 2002
Rwanda	1991, 2002
Saint Lucia	1980, 1991
Senegal	1988, 2002
Sierra Leone	2004
Spain	1981, 1991, 2001
South Africa	1996, 2001, 2007
South Sudan	2008
Sudan	2008
Switzerland	1970, 1980, 1990, 2000
Tanzania	1988, 2002
Thailand	1970, 1980, 1990, 2000
Turkey	1985, 1990, 2000
Uganda	1991, 2002
United Kindom	1991
United States	1960, 1970, 1980, 1990, 2000, 2005, 2010
Uruguay	1963, 1975, 1985, 1996, 2006
Venezuela	1971, 1981, 1990, 2001
Vietnam	1989, 1999, 2009