CAMPINAS STATE UNIVERSITY
INSTITUTE OF PHILOSOPHY AND HUMAN SCIENCES

GUSTAVO PEDROSO DE LIMA BRUSSE

"CONTRIBUTIONS OF HOUSEHOLD PROJECTION METHODS IN THE BRAZILIAN CONTEXT: A COMPARATIVE STUDY FOR SÃO PAULO STATE"

"CONTRIBUIÇÕES DE MÉTODOS DE PROJEÇÕES DE DOMICÍLIOS NO CONTEXTO BRASILEIRO: UM ESTUDO COMPARATIVO APLICADO À POPULAÇÃO DO ESTADO DE SÃO PAULO"

CAMPINAS
2021

GUSTAVO PEDROSO DE LIMA BRUSSE

“CONTRIBUTIONS OF HOUSEHOLD PROJECTION METHODS IN THE BRAZILIAN CONTEXT: A COMPARATIVE STUDY FOR SÃO PAULO STATE”

“CONTRIBUIÇÕES DE MÉTODOS DE PROJEÇÕES DE DOMICÍLIOS NO CONTEXTO BRASILEIRO: UM ESTUDO COMPARATIVO APLICADO À POPULAÇÃO DO ESTADO DE SÃO PAULO”

Thesis presented to the Institute of Philosophy and Human Sciences of the Campinas State University in partial fulfillment of the requirements for the degree of Doctor in the area of Demography.

Tese apresentada ao Instituto de Filosofia e Ciências Humanas da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Doutor na área de Demografia.

Orientadora: PROFA. DOC. TIRZA AIDAR

THIS WORK CORRESPONDS TO THE FINAL VERSION OF THE THESIS DEFENDED BY STUDENT GUSTAVO PEDROSO DE LIMA BRUSSE AND SUPERVISED BY PROF. DOC. TIRZA AIDAR.

CAMPINAS
2021

FICHA CATALOGRÁFICA

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Filosofia e Ciências Humanas
Paulo Roberto de Oliveira - CRB 8/6272

B838c

Brusse, Gustavo Pedroso de Lima, 1989-
"Contributions of household projection method in the Brazilian context : a comparative study for São Paulo state" / Gustavo Pedroso de Lima Brusse. – Campinas, SP : [s.n.], 2021.

Orientador: Tirza Aidar.
Coorientador: Glaucia dos Santos Marcondes.
Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Filosofia e Ciências Humanas.

1. Método de projeção. 2. População. 3. Domicílios. I. Aidar, Tirza, 1961-. II. Marcondes, Glaucia dos Santos, 1974-. III. Universidade Estadual de Campinas. Instituto de Filosofia e Ciências Humanas. IV. Título.

Informações para Biblioteca Digital

Título em outro idioma: "Contribuições de métodos de projeções de domicílios no contexto Brasileiro : um estudo comparativo aplicado à população do estado de São Paulo"
Palavras-chave em inglês:
Projection method
Population
Households
Área de concentração: Demografia
Titulação: Doutor em Demografia
Banca examinadora:
Tirza Aidar [Orientador]
Brenda Yepéz Martínez
Gustavo Henrique Naves Givisiez
Joice Melo Vieira
Marcos Roberto Gonzaga
Data de defesa: 11-06-2021
Programa de Pós-Graduação: Demografia

Identificação e informações acadêmicas do(a) aluno(a)
- ORCID do autor: https://orcid.org/0000-0002-5506-3610
- Currículo Lattes do autor: http://lattes.cnpq.br/0541850429767181

CAMPINAS STATE UNIVERSITY

INSTITUTE OF PHILOSOPHY AND HUMAN SCIENCES

A comissão julgadora dos trabalhos de Defesa de Tese de Doutorado, composta pelos Professores Doutores a seguir descritos, em sessão pública realizada em 11 de junho de 2021, considerou o candidato Gustavo Pedroso de Lima Brusse APROVADO.

Profa. Dra. Tirza Aidar – Orientadora

Prof. Dra. Brenda Yepéz Martínez – Universidad Central de Venezuela

Prof. Dr. Gustavo Henrique Naves Givisiez – Universidade Federal Fluminense

Prof. Dr. Marcos Roberto Gonzaga – Universidade Federal do Rio Grande do Norte

Prof. Dra. Joice Melo Vieira – IFCH/UNCIAMP

A Ata de Defesa com as respectivas assinaturas dos membros encontra-se no SIGA/Sistema de Fluxo de Dissertações/Teses e na Secretaria do Programa de Pós-Graduação em Demografia do Instituto de Filosofia e Ciências Humanas.

To my lovely family: Dino, Elsa, Malu, Brusse, Felipe, Gabi, and Spock.

ACKNOWLEDGMENTS

Firstly, I want to express my gratitude for my supervisor Tirza Aidar for your support and advice, for giving me the opportunity to work on such interesting field and having patience to transmitting all your experience. I believe we have made a good team working together during all these years. Also, I want to thank my co-supervisor Glaucia dos Santos Marcondes for her support and her broader view that gave consistency and focus to the thesis.

To Joice Melo Vieira who encouraged me to be a demographer since the scientific initiation during my undergraduate studies in statistics and to Brenda Yépez Martínes whom work inspires me and served as a reference. Both were members of my thesis's qualification exam and contributed to improving this thesis.

To all professors from Demography department at UNICAMP who contributed to my graduation formation: Luciana, Zé Marcos, Maria Coleta, Rosana, Everton, Roberto, Roberta, Tirza, Glaucia, Joice, Ana Silvia, Marta, Bete Bilac and Alberto. A special thanks to all academic and technical support staff in Demography department and Núcleo de Estudos da População Elza Berquó (NEPO): Bene, Betanho, Leandro, Marcelo, Adriana, Rogério, Raquel, Denise, and all cleaning and I.T. support staff.

Many steps were fundamental to the development of this thesis. In 2017, the seminar "Proyecciones de hogares: desafíos conceptuales y metodológicas para las poblaciones de América Latina" hosted by Núcleo de Estudos da População Elza Berquó (NEPO) was an importante space to discuss household projection in Latin America with presence of great researchers in the field, such as, Brenda Yépez Martínez, Maria Coleta de Oliveira, Elisabete Bilac, Tirza Aidar, Guiomar Bay, Roberto Luiz do Carmo, Landy Lizbeth Sanchez Peña, Eduardo Torres, Paulo Jannuzzi, Bernadette Cunha Waldvogel and Carlos Eugenio de Carvalho Ferreira.

Another essential step was the Household Projection Workshop entitled "Population and household forecasts with POPGROUP" during the first months of 2018 conducted by Professor Ludi Simpson (University of Manchester), when I first had contact with the Headship Rate method.

The wonderful time spent in European Doctoral School of Demography (EDSD) held in Max Planck Institute for Demographic Research (MPIDR) in Rostock/Germany and the University of Southern Denmark (SDU) in Odense/Denmark during 2018/2019. Thanks to my EDSD mentor Jim Oeppen and all EDSD teachers: Jim Vaupel, Virginia Zarulli, José

Manuel Aburto, Alyson van Raalte, Tim Riffe, Mircea Trandafir, Daniele Vignoli, Christian Dudel, Giancarlo Camarda, Jutta Gampe, Jim Carey, Francesco Villavicencio, Jesus Álvares, Annette Baudisch, Ilya Kashnitsky, Ugofilippo Basellini, Vladimir Canudas Romo, Albert Esteve, Silvia Rizzi. To my lovely classmates Sarah, Alessandro, Francesco, Rae, Jiaxin, Marcus, Raphael, Martina, Isabelle, Marie, Cecilie, Ricarda, Steffen, Orsola and all roommates at dorms. Special thanks to Marie-Pier Bergeron and Heather Booth for teaching forecast techniques some of them used in this thesis.

I gratefully acknowledge Prof. Yi Zeng (Duke University, Peking University) and Prof. Zhenglian Wang (China Population and Development Research Center) for teaching the method they have developed for years and for valuables suggestions and discussions alongside during this the research project and during the "International Conference and Training Workshop on Household and Living Arrangement Projections for Informed Decision-Making" held in Beijing, China in April 2019.

To all my friends and classmates from UNICAMP, Ana Camila, Andrews, Karol, Blad, Camila, Caio, Chandeline, Cimar, Dafne, Dannyra, Dario, Dayane, Felipe, Fernanda, Giovana, Guilherme, Ianca, Igor, Jaque, Jóice, Joyce, Juliana, Késia, Katiane, Kelly, Leandro, Letícia, Leonardo, Livia, Livan, Luiza, Luiz, Marcelo, Marina, Manu, Milena, Nathália, Natália, Pedro, Raíssa, Raphael, Rafa, Renato, Rignaldo, Ricardo, Thiago, William, Sofia, Thati, Vinicius, José Henrique.

I would like to thank SEADE Foundation, especially Bernadette Cunha Waldvogel, for their assistance with data collection.

Finally, I want to thank all my family Dino, Elsa, Malu, Brusse, Felipe, Gabi, Spock and my closest friends for their support and encouragement throughout my studying years and for whom I dedicate this work.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001. And part by Netherlands Interdisciplinary Demographic Institute (NIDI).

ABSTRACT

Household projection is especially important among studies concerning social housing demand, family formation and environmental issues such as consumption of water, energy, and durable goods. Despite some few attempts, Brazil does not have an official practice of household projection at the national or subnational level and those previous attempts are mostly restricted to the total number of households, not taking into account age and sex distribution, neither the size nor composition of these future households. Taking São Paulo state (Brazil) as the base population, this study compared three household projections: one using the Cohort-Component Method combined with Headship Rate Method, implemented in PopGroup software; another one using the Extended Cohort-Component Method, developed by Zeng et al. (1991) and implemented in ProFamy software; and finally, one alternative approach using a Cohort-Component Method implemented in R with Functional Principal Components (FPC) models to project mortality and fertility combined with a new method for Headship Rate forecasting. Each method represents a different type of static and dynamic household projection approach. They have distinct technical characteristics such as different inputs variables and assumptions. Five scenarios were created combining high/low mortality, high/low fertility, and constant rates, while for the new approach 95% confidence intervals were used to create new scenarios. The results were compared in terms of number of households with the 2010 Census observed data (projection from 2000 to 2010) and the projections made by the SEADE Foundation (Official São Paulo State Statistic Office) for 2050. The main goal of this work consists in contribute to methodological improvements in household projections, that means more quality information available to demanding fields and justify the systematic adoption of household projections in Brazil.

Keywords: Projection method; Population; Households.

RESUMO

As projeções de domicílios são importantes, não apenas para entender a complexa relação entre população e domicílio, mas também possuem aplicações em áreas como a demanda demográfica por domicílios (mercado imobiliário, habitação social, planos diretores municipais), consumo e meio ambiente (demanda do uso de água, esgoto, energia elétrica, bens duráveis, uso de automóveis) e os próprios estudos da área de Família (cuidado de idosos e crianças, disponibilidade de parentes, relações intergeracionais). No Brasil, não há uma prática oficial de projeções domiciliares a nível nacional ou subnacional e as poucas experiências anteriores são restritas ao número total de domicílios, não levando em consideração a distribuição por idade e sexo, nem o tamanho e a composição desses futuros domicílios. Utilizando o estado de São Paulo como população base, esse estudo compara três diferentes métodos de projeção de domicílios possíveis de serem aplicados com fontes de dados brasileiros: o primeiro utilizando o tradicional Método de Coortes Componentes combinado com o Método das Taxas de Chefia, implementado no software PopGroup; o segundo utilizando o Método Estendido de Coortes Componentes, implementado no software ProFamy; e um método alternativo proposto que utiliza o Método de Coortes Componentes implementado no software R, baseado em modelos funcionais para projeção de mortalidade e fecundidade e uma nova abordagem para projeção das Taxas de Chefia, baseado em Lee e Carter (1992). Com isso, espera-se contribuir com refinamentos metodológicos de projeções de domicílios que avancem na qualidade de informação disponibilizada para áreas que demandam este tipo de informação e que justifique a adoção sistemática de projeções domiciliares no Brasil.

Palavras-chave: Método de Projeção; População; Domicílios.

FIGURES LIST

TABLES LIST

SUMMARY

INTRODUCTION

The relation between population and households is not straightforward. Even if a given population growth is low or negative over the time, the number of households can still grow depending on the living arrangements distribution and the decreasing household size of that population. The increasing time spent in the parents' home, the decreasing number of children, changes in marital status pattern and the increasing longevity are some of the demographic life course events that can modify the number, size, and composition of future households of a population. Not only to understand better this complex relationship, but household projection models are also important among studies concerning social housing demand, family formation and environmental studies such as consumption of water, energy, and durable goods.

In social housing studies, for example, policy makers are interested in the flow of housing deficit and stock measures in order to plan how many housing units will be necessary to supply the social housing demand of a certain region at certain period (GIVISIEZ; OLIVEIRA, 2018; ZENG et al., 2014a). However, the need for new housing units resulting from the demographic dynamics can vary by age and sex. While young adults are exposed to a period of life when the formation and dissolution of households is more intense, population aging increasingly places more importance on the elderly households. In Hebei, China, Zeng et al. (2014a) show that compared to 2010, the housing demand of those aged 65 or more will increase by 228.3% and 447.6% in 2030 and 2050. In addition to the number of households by age and sex, the size of these housing units is especially important to optimize the governmental resources allocation in social housing policies (GIVISIEZ; OLIVEIRA, 2018).

Other applications where household projections are considerable relevant is for family studies. In ageing societies, not only the future number of the elderly will matter, but by whom this growing population will be cared for. The elderly care supply is strongly associated with co-residence with spouse and relatives of other generations (ex. children and grandchildren), or with relatives of the same generation (as siblings) who would provide care inside the household. In east Asian countries, such as Fiji, Republic of Korea, Malaysia and Philippines, elderly co-residence with other generations is about 75% and 85% of total elderly care (SOKOLOWSKY, 2008). In developed countries, 1/3 of elderly care is carried out in institutions, while the other 2/3 are carried out within the households (NERI; SOMMERHALDER, 2002). The family is the institution that is responsible for most of the elderly and child care and household projections can give essential information about future kinship availability that will determine the care demand and care supply (CAMARANO; KANSO, 2010).

The future size and composition of households is also important to environmental and consumption studies since the consumption pattern is different between cohorts and it can change over the lifetime of the individuals. Consequently, the participation percentage of certain age groups inside households will have a direct impact on the consumption curve of a population (CASTILLO; PEÑA; GUARDIÁN, 2016; SILVA, 2013). Furthermore, the reduction in household average size has a double impact on the use of resources and on biodiversity (LIU et al., 2003). First, there is an increase in the demand for land and materials (e.g., wood, concrete and steel) needed for housing construction. Secondly, households with fewer residents have less efficiency in per capita resource usage, because goods and services are shared by more people in larger households, such as some fixed environmental costs of energy and water, production and deposition of waste and sewage (LENZEN; MURRAY, 2001; IRONMONGER; AITKEN; ERBAS, 1995; YOUSIF, 1995). Also, some studies verify that the number of cars, consequently the emissions of carbon monoxide and hydrocarbons by cars, are associated with the age of household head and the size of households (PRSKAWETZ; LEIWEN; O'NEILL, 2004; FIORAVANTE, 2009). In Brazil, more than 4.6 million households were created between 1991 and 2000 due to a reduction in average household size. The contribution of the reduction of the average size in the total number of households is 46% while 54% is due to population growth (LIU et al., 2003). Between 1991 and 2000, the population grew at an annual rate of 1.7% (from 145.6 to 168.3 million inhabitants), while households increased by 3.2% (from 34.7 million to 44.7 million households).

All these research areas demonstrate that household projections must move toward more detail outcomes which take into account not only the total number, but the age, sex, size and composition distribution of these future households. But the question is, which demographic, mathematical or statistical model is able to express the complex relation between population and households and produce such detailed information of the future?

Since the beginning of the 1980s, Family and Household Demography has grown rapidly in the international scientific community with the publications of Burch (1979); Bongaarts (1983); Brass (1983); Ryder (1985); Bongaarts; Burch and Watcher (1987); De Vos and Palloni (1989) and Grebenik and Mackensen (1989). Four types of models for projecting households have been developed: analytical models, macrosimulation, microsimulation and probabilistic models. The analytical models aim to establish relations between stable population parameters and kin availability, while in the microsimulation models, the individuals are the "units of analysis and transition rates are applied to each person from an initial population, generating individual trajectories and thus a whole distribution" (WAJNMAN, 2012, p. 33). Macrosimulations, on the other hand, "apply transition rates to groups of aggregate individuals, the cohorts, producing trajectories representing average individuals in the cohorts" (WAJNMAN, 2012, p. 33). Also, some software's have been developed to implement those models, such as SOCSIM (HAMMEL et al., 1981), LIPRO (VAN IMHOFF; KEILMAN, 1991), KINSIM (WOLF, 1990), MOMSIM (RUGGLES, 1993), STINMOD and DYNAMOD from Canberra National Centre for Social and Economic Modelling (NATSEM), MOSART-H from Central Bureau of Statistics of Norway and NEDYMAS from Netherlands Central Bureau of Statistics.

Finally, more recently probabilistic models try to incorporate variability and estimate confidence intervals for projections. All these types of household projections models can also be classified as *dynamic*, when the cohort behavior over time is studied, or *static*, when the model studies the future distribution of households based only at a period of time. This study will focus on the *dynamic* and *static* model definitions.

Yepez Martínez (2010) reviewed the household projections of 68 National Statistical Offices. The experiences were compared on method used (dynamic or static), household type classification, data sources, geographic area and projection year. The author found out that most countries still use static methods for projecting households. The models developed after the 1980's Family Demography boom have not yet been absorbed by National Statistical Offices, which still adopt the United Nations recommendation that place the best-known static method (the Headship Rate Method), as the most appropriate method for household projection and which had advantages over the other methods available at the time (UNITED NATIONS, 1973). According to Imhoff et al. (1995) the requirement of longitudinal data is the main impediment for dynamic methods to be used since most of the countries still remain in a context of cross-sectional data from censuses and surveys. Also, the investment in paid software and the complexity of these methods which requires a larger set of inputs and calculations are some of the barriers for national and subnational official statistical offices to take a step toward dynamic methods.

In Brazil, some studies concerning household projections have been done in academic studies, agreements with public administration institutions and consultancies for energy companies, all of them using static methods (LEON; MOREIRA, 2005; EPE, 2004; ELECTROBRÁS, 2007; GIVISIEZ; OLIVEIRA, 2011; CEDEPLAR, 2007; SEADE

Foundation, 2017). Despite these few attempts, Brazil does not have an official practice of household projection at the national or subnational level and those attempts are restricted to the total number of future households, not considering the size nor the composition of these households. Also, most of those methods use outdated mortality and fertility forecast models, based on private software which the user cannot control the projection process, manipulate the dataset, or plot their own results. Often, they use flawed scenarios assumptions such as constant rates.

In this study, three household projection methods will be applied. Two of them represent a different type of static and dynamic approach and their results refer to the total number, size, and composition of households. The last method is an alternative approach aiming to improve the classical Headship Rate Method in an open-source software. The central purpose here is to evaluate whether Brazil has enough data source availability and data quality to support a step toward a dynamic approach compared to the classic static approach, and how can we improve the classical static approach to reach better results in order to contribute to a regular practice of household projections in Brazilian national and subnational statistical offices.

All data and codes will be available in https://github.com/GustavoBrusse. Due to text organization, graphs with different scales may be presented together. For comparisons, please always check the scale of the graphs.

CHAPTER 1 – HOUSEHOLD PROJECTION METHODS STUDIES: BACKGROUND AND SIGNIFICANCE

1.1 From individual to household and family dynamics

The demographic dynamics can be understood as a study of changes in the number of individuals for a given population over a period, resulting from the balance between the three fundamental components of populational change and the respective relationship they establish among themselves: mortality, fertility, and migration.

In a similar way, the demographic dynamics analysis can be applied in groups of individuals such as families and households. The family and household dynamics can be understood as the study of its components change and not rare is called “Family and Household Demography” in the international literature. In this way, the interest of this field refers to the study of “mortality” (i.e., the dissolution of families or households), the study of “birth” (i.e., the formation of families or households), as well as to the change of its composition, mainly in terms of the household types that prevail in a specific population and the age and sex composition of its residents.

Unlike ‘individuals’ demographic dynamics, in which death and birth processes are unique and well-defined events, the family and household formation and dissolution events are more complex. They involve not only fertility (number of children who are co-residents) and mortality (co-resident’s deaths), but also depend on the nuptiality dynamics (marriages, unions, separations, and divorces) and the departure/return from parents’ home, events that can be repeated in the individual’s life course and have more abstract definition than essentially biological events.

Although family and households are often referred to as the most important mediation level between individual and population, the relationship between individual based-demography and family and household-based demography is not straightforward. For example, even in a population whose population growth rate is low or even negative, the number of households may continue to grow depending on the household distribution and the average household size for that population (ZENG et al., 2014b; GU et al., 2015).

Most of the countries find themselves in this stage: have experienced the transition from high fertility and high mortality rates regime to low fertility and low mortality rates regime, resulting low population growth rates. In that scenario, the literature has given more and more attention to the relationship between demographic dynamics and family and household dynamics. As example of that, one of the main arguments that support a supposed

Second Demographic Transition (SDT) theory is the existence of a family structure change (LESTHAEGHE; VAN DE KAA, 1986). According to this, it would be observed in post-demographic transition stage countries, a systematic drop in marriage's proportion and marriage rates, an increase in the average age at first marriage and separations and divorces occurring more frequently and in earlier ages.

However, the attempt to identify modification of universal patterns of family and household structure is not particular a demographic concern, nor it is a novelty in the literature. Tocqueville (1969); August Comte, Le Play (1871); Durkheim (1888); Westermarck (1891) and more recently Parsons (1949); Goode (1959) and Therborn (2004) were some of the authors considered as classics that dealt with the theme. Several theories on the families and household behavior trends have been discussed in areas such as Sociology, Anthropology, Economics and History. Among them, there is the hypothesis that there would be a global convergence towards the nuclearization of the family and households (predominance of families composed only of couples with reduced numbers of children) and the hypothesis that the composition will be in constant flux, alternating times of greater arrangements diversity for periods of predominance of family nuclearization. Such assumptions often vary from the perspective of the family's resistance as a solid institution, to the fluidity and heterogeneity of the household types.

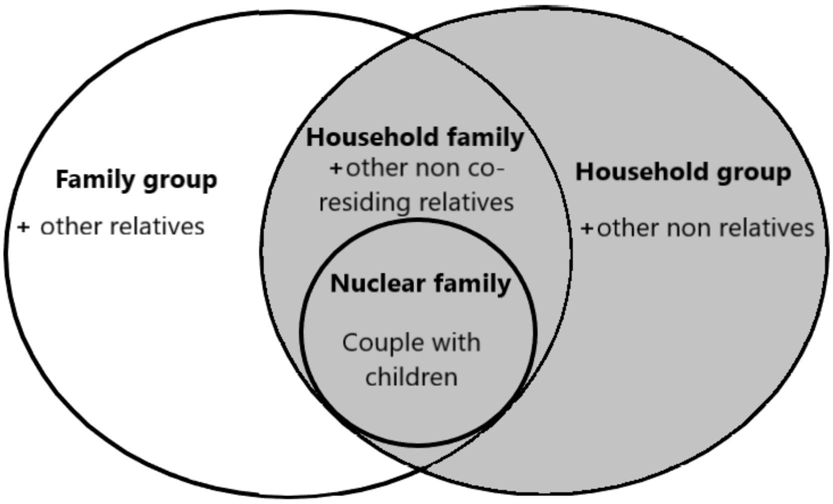
Although repeatedly discussed, these hypotheses often fail to state with certainty whether there is a tendency for families and household change over time, most of the time because of the lack of empirical evidence for these findings. In a recent effort, the United Nations published the "Household Size and Composition Around the World 2017" (UNITED NATIONS, 2018), which unifies and harmonizes data on composition and household types for 163 countries using 745 data sources including censuses and surveys. It makes a worldwide review of the recent trend of what is happening in the households and concludes that the only true generalizable conclusion is that there is a drop in the average household size, due to a decline in extended households and an increase in living alone households. According to this study, nothing more can be said in terms of trend.

In fact, as Yepez Martínez (2010) points out, changes in the household composition cannot be explained only by Family theories. It must also be explained by the demographic changes that have occurred in the population and that consequently will have an impact both structural and organizational dimensions of households. Many of the observed phenomena, such as the decrease in the average household size, have a strong demographic influence and not just a change in family or individual choices.

In this sense, Demography offers a range of different concepts framework and methodologies compared to Economics, Sociology, Anthropology and History. Some authors argue that, what the Family Demography has as a particular characteristic is “the empirical treatment of information, through a set of methods that allow to synthesize and model the processes, in order to predict future behaviors based on hypotheses about its components” (WAJNMAN, 2012, p. 17).

As a fundamental tool for family and household empirical measurement, Demography uses the household definition and the relation to the reference person question present in several data sources’ questionnaires. As defined in Wajnman (2012), the family can be understood as a group that involves a net of individuals connected by consanguinity and/or alliance who may or may not live in the same household. During this work, the terminology “household” will be used to refer to the “household group” concept, that is, that population that has the common bond of living in the same household unit, being a relative or not (group hatched in the Figure 1).

FIGURE 1 – Concept diagram of “Family group”, “Household group” and “Household family”



Source: Wajnman (2012).

According to Cavenaghi and Alves (2011) the household definition used by Brazilian censuses (except in 1960 and 1970) is equivalent to the concepts of “Vivenda” in Spanish or “Housing units” in English, and follows the UN recommendation, which defines a Housing unit as: “a separate and independent place of abode intended for habitation by one household” (UNITED NATIONS, 2007, p. 238). The concept of “Hogar” (Spanish) or “Household” (English) has never been used in Brazilian census surveys, and despite of the existence of a specific word in the Brazilian vocabulary that expresses the concept called "fogus", the term "domestic unit" is used by IBGE as the closest expression for its translation. According to the UN definition, a domestic unit, "Hogar" or "Household" is a unit in which people, or one person, form a group to satisfy their needs for food and other essential items for living (UNITED NATIONS, 1998; 2007).

It is also important to mention that a household can be divided into three categories in Brazilian census: "permanent and private", "unoccupied" and "collective". The "permanent and private" households are built and occupy exclusively to serve as housing purposes and represent a major part of the total households. This category can be projected through the demographic projection methods and will be our object of study, since demographic information are known from their residents (mortality, fertility, and migration). The "unoccupied" households are divided into three other categories: "closed at the time of the census", "occasional use" and "vacant". It includes, respectively, where it was not possible to perform the census interview, where it was being used for weekend breaks, vacations, or other seasonal purposes, and where the unit was available for sale or rent. As there is no population living in them, it is unable to be demographically projected. While "collective" households, which represents less than 1% of Brazilian households, are long-term institutions, prisons, military quarters, camping sites, hospitals, hotels, and other establishments where residents live under the rules of an administrative subordination. The total number of households observed in a given region is the sum of "permanent and private", "unoccupied" and "collective" households. Also, the "permanent and private" households can be classified as "improvised" when it is located in a building (store, factory, building under construction, tent, wagon, etc.) that did not have a dependency dedicated exclusively to housing, as well as an inappropriate place for housing, which, on the reference date, was occupied by a resident.

The question about the household's reference person is traditionally used in census questionnaire to identify individuals within the same household and to know the relationship between each member in that household. First, the interviewee chooses an individual from that household as the reference person. From that, a list of residents and their respective relationships (kinship or not) with the reference person is made. Over time, there have been changes in terminology regarding the reference person definition, as well as the number of categories to define the relationship with the reference person. The consequences of that changes will be further discussed.

From this information, multiple household typologies can be done. The household typology chosen here takes into consideration the international comparability and the comparability between different Brazilian data sources and can be written as:

- **Living alone** (one person living alone, excluding people living in collective households who are also considered reference persons in some Brazilian censuses)
- **Couple without children** (reference person and spouse)
- **Mono-parental** (reference person and children)
- **Couple with children** (reference person, spouse, and children)
- **Extended household** (any type of previous household that additionally contains “other relatives” and/or “non-relatives”)

Table 1 and 2 present the household type distribution in some developed and developing countries. One condition that distinguish Brazil and other Latin American and developing countries from developed countries is the maintenance of a high percentage of extended households around 20% to 25% of total households, high percentage of couple with children and lower percentages of people living alone and couple without children.

TABLE 1 – Household type distribution (%) in some developed countries according to the last Census information on IPUMS

Household type	USA	UK	Norway	Japan	Italy	Spain	Australia	Germany
Living alone	27.89	30.58	39.58	34.53	30.91	23.19	22.83	37.27
Couple with children	23.82	25.21	27.08	26,79	33.76	34.96	29.45	23.32
Couple without children	24.79	25.56	21.41	20.10	18.85	21.04	24.39	28.39
Mono-parental	9.13	10.69	8.38	8.90	9.41	9.36	9.93	7.33
Extended household	9.02	N.I.	N.I.	N.I.	6.28	N.I.	N.I.	N.I.

Source: United Nations (2018).
N.I. – No information

TABLE 2 – Household type distribution (%) in some developing countries according to the last Census information on IPUMS

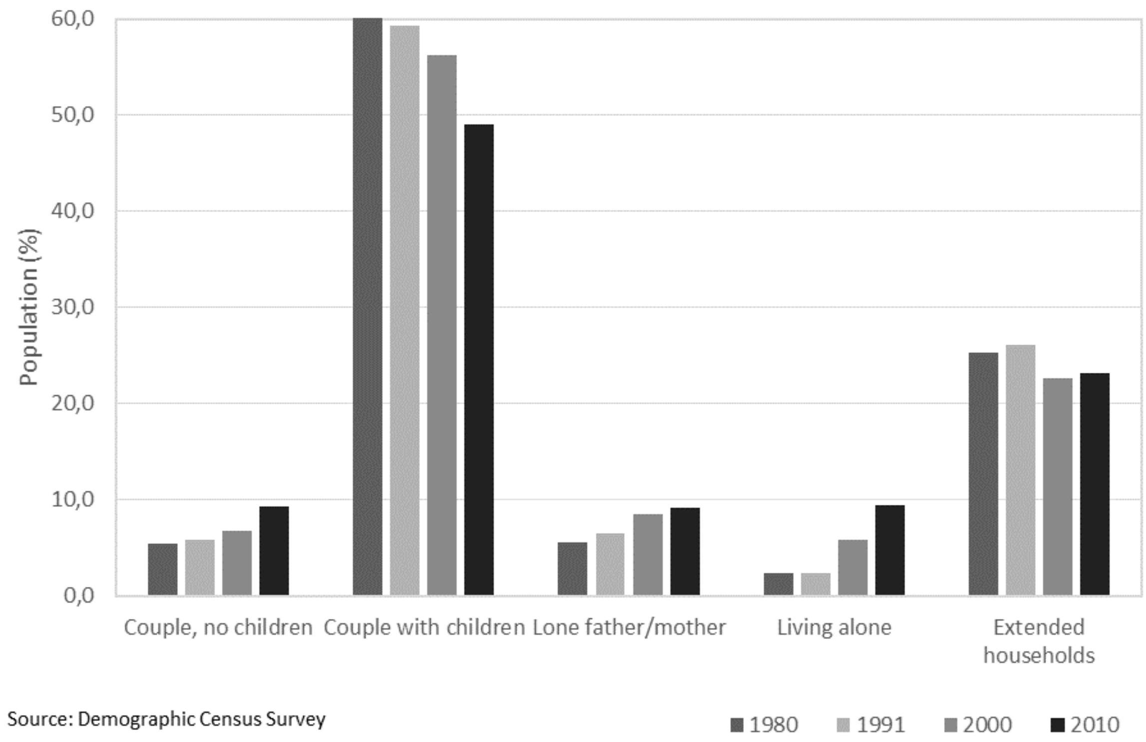
Household type	Brazil	Russia	China	South Africa	Mexico	Argentina	Turkey
Living alone	12.00	25.97	8.59	23.84	10.8	17.60	6.37
Couple with children	40.41	23.57	46.96	15.60	41.60	36.77	55.08
Couple without children	13.90	15.12	12.82	8.70	10.2	13.34	13.33
Mono-parental	10.47	11.66	5.40	11.56	10.16	11.73	4.59
Extended household	20.86	22.61	N.I.	37.70	24.72	18.07	19.46

Source: United Nations (2018).
N.I. – No information

Extended household is also found in literature named as “cohabiting family household”, “working class family household”, “patriarchal family household” or “poor family household”. Co-residing with other relatives beyond nuclear family is usually associate with a cultural behavior and an economic strategy in period of economic crises and high unemployment. The 2012 Brazilian National Household Survey (PNAD) asked about what the motivations were to people shared the household with relatives and a pull of necessities were cited, such as, financial problems, teenager pregnancy, elderly health care, later insertion in labor market, and additional years of education. At the same time, Ruggles (2007) shows the decline of intergenerational coresidence from 1850 to 2000 in the United States with a powerful association between higher opportunity for the younger generation and low intergenerational coresidence. Consequently, all these reasons are also related to differences in living alone among developing and developed countries.

From reference person and household type definitions, it has been possible to follow the household type growth trend by demographic censuses and PNADs for at least 30 years. With these definitions, it is also possible to analyze data from two different perspectives: the perspective of population living in households, or the perspective of number of households in the population. The graph below shows the population proportion of São Paulo’s state according to the household typology above in the last censuses. That is one way of looking to what has been happening in the São Paulo’s state in terms of household trends for the total population.

FIGURE 2 – São Paulo's total population by household type, Census 1980, 1991, 2000 and 2010



Over the analyzed time, there were no great changes between population proportions in household types in terms of the São Paulo's total population. However, it is possible to see more substantial drop in the population proportion number living in more "conventional" couple with children type, which went from approximately 60% of the state's population in 1980 to just under 50% in 2010. Meanwhile, there was an increase in population living in "living alone", "couple without children" and "mono-parental" households, and a fluctuation around the 5% of population living in extended households.

Bennett and Dixon (2006) show some trends regarding the living alone household formation. First, there has been a marked increase in the number of young adults living alone, particularly between the ages of 25 and 44. In addition, the proportion of people living alone in such age group is higher for men than for women. This situation is reversed with advancing age. For people over 75, the proportion of those living alone is much higher for women, given the greater female longevity and due to the greater number of male remarriages.

Regarding mono-parental households, Oliveira; Sabóia and Soares (2002) argue that, although there is a convergence process in relation to the headship rates of these household by sex, female heads are the vast majority in mono-parental arrangements and have a strong urban character. Therefore, the growth of this household type is related to young women in large cities.

In Brazil, as well as in most developed and developing countries, changes in population proportions by household types are mainly linked to the decrease in the number of children born, the increase in longevity, the longer time spent at the parents' home and changes in nuptiality, as highlighted by Aidar et al. (2017):

> The impressive drop in fertility, which rapidly reduces the young-age-dependency ratio and, together with the constant decrease in mortality over the last twenty years, results in a process of population aging; the changes of timing in the markers of the different phases of the life course; the transformations in gender relations, the greater visibility of homosexual unions; the increase of consensual unions and divorces in the last decades, *pari passu* the increase of formal marriage rate due to remarriage of the divorced and collective wedding promoted by public services and churches (AIDAR et al., 2017, p. 1).

Some authors summarize the impacts of changing fertility and mortality rates on the number, size and composition of families and households as the "verticalization" process of families (KNIPSCHEER, 1988; BENGTSON; ROSENTHAL; BURTON, 1990; GEORGE; GOLD, 1991; CAMARANO; KANSO, 2010). In this process, families and households lose relatives of the same generation while there is an increase co-residence between different generations. Successive cohorts that experienced low fertility rates, reduces the household size with decline number of sons and daughters, consequently the number of siblings, cousins, uncles, and aunts. While reduced mortality rates at advanced ages increase the co-residence of parents, grandparents, and great-grandparents. Murphy (2010) verified with England data that in 1950 a child under 5 years old had an average of 2.5 living grandparents. In 2000, this average rose to 3.5 living grandparents and in 2010 a child could spend his entire childhood living with 3 living grandparents on average. However, while grandchildren have more grandparents alive, grandparents have fewer grandchildren due to the fall in fertility and the postponement of the average age of birth of the first grandchild (MURPHY, 2010). The average age of the first grandchild was 55 years old in 2000 and will be 70 years old in 2050 in England (MURPHY, 2010).

In addition to "verticalization" process, the departure/return from the parents' home is also determinant for new household formation and dissolution at certain ages. The literature calls attention to the "kangaroo generation": young people who have a postponement of leaving their parents' home in transition to adulthood phase of life (MODELL; FURSTENBERG JR.; STRONG, 1978; GOLDSCHEIDER; DAVANZO, 1986; SETTERSTEN; FURSTENBERG; RUMBAUT, 2008; VIEIRA, 2009). Because of that, we expected a household dissolution postponement of the "couple with children" household type and a formation postponement of the "living alone" and "couple without children" household types. Meanwhile, parents will experience older the so-called “empty nest” phase of life. In Brazil, the proportion of young people between 25 to 34 years of age who still lived in their parents’ homes raised from 21.7% in 2005 to 25.3% 2015 which 60.2% of them were men (AIDAR et al., 2017).

Combined to the verticalization process, the later departure from the parents’ home is also decisive for “sandwich generation” formation, the middle-aged mothers who care for both their young children and the elderly parents with whom they co-reside with (GOLDSTEIN; MASON; ZAGHENI, 2011; DUKHOVNOV; ZAGHENI, 2015, LIMA; TOMÁS; QUEIROZ, 2015). Despite of increasing number of “kangaroo generation”, Tomás; Lima and Queiroz (2018) show that the “sandwich generation” is shrinking in the last decades in Brazil.

Leaving the parents’ home is often followed by marital status transition. Therefore, the postponement of leaving the parents’ home is also associated with an increase of average age of first marriage and average age of first union (VIEIRA; ALVES, 2016). For São Paulo’ state, the male average age at first marriage in 2000 was 31.5 years old while the average age at first union was 26 years old (VIEIRA; ALVES, 2016). For women, the average age at first marriage in 2000 was 28.8 years old while the average age at first union was 23 years old (VIEIRA; ALVES, 2016). Yet, the unions and marriages dissolution can represent a return to the parents’ home or an increase in living alone and mono-parental households among adults. Between 1984 and 1990 the number of divorces more than doubled in São Paulo state (FREITAS, 2019).

Migration as a component of population change also affect household formation and dissolution. Previously well-known, migration often happens with the family and young people during important transitions phases. In a context where natural growth rates are getting lower and lower, migration plays an increasingly important role to population change. Meng and Gregory (2005) and Duncan and Trejo (2007) argue that immigrants tend to marry other immigrants living in the country of residence or to marry immigrants who migrate for the purpose of marriage. Therefore, the migration of entire families as well as individuals can significantly change the household distribution of a population, establishing migration neighborhoods or regions that may have a different household pattern compared to the native population.

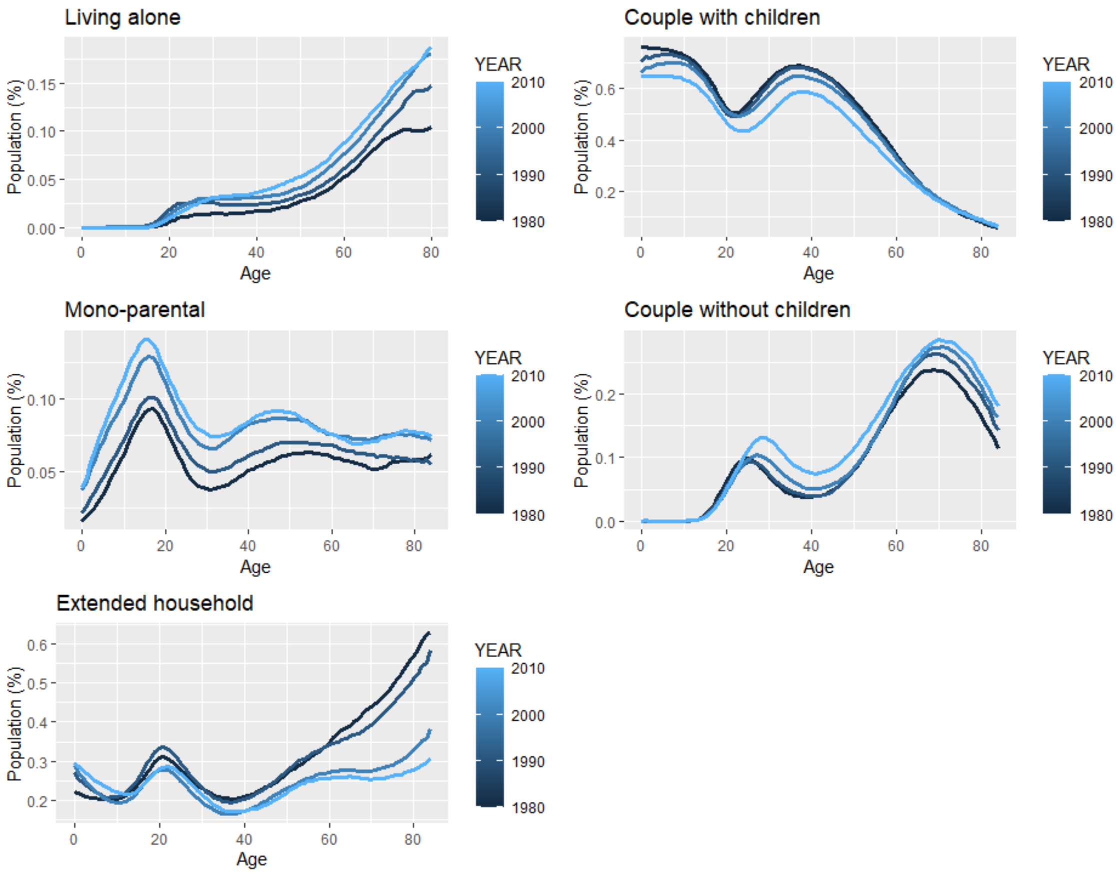
All these discussed phenomena, from mortality and fertility, to leaving the parents’ home, nuptiality and migration have important differentials among geographic regions and educational, socioeconomic and race/color/ethnicity groups. As a consequence of different patterns of household formation and dissolution processes, there will be different future household number, size and composition trends for each population group. This is even more evident in the Brazilian society, historically marked by deep inequalities (BILAC, 2017).

In this sense, the age variable is an essential tool in demographic analysis that helps to highlight different processes of household formation and dissolution of among population subgroups. All these changes discussed earlier get a temporal approach looking from age perspective and it allows us to see family and individuals' lives along their own trajectory. Two methodological approaches are possible from this: the family life cycle approach (GLICK, 1977), that studies how families and households contract or expand throughout the family lifetime, also analysis the reduction of their average size; and the life course approach (ELDER, 1978; HAREVEN, 1994) that studies the individuals or population aggregates trajectory of over time, enabling the identification of age markers and transitions between different states and turning points. The Life Course approach differs from the classic approach that considers only a fixed moment in time, known as "cross-sectional" (GOLDANI, 1989).

Looking back to the recent household changes in São Paulo's state (showed in Graph 1) from the age perspective, it is possible to see not only the change in population proportion residing in each household type over time, but also important ages "markers" transitions, postponing or delaying certain types of household formation or dissolution (Graph 2). An example can be seen at young adult population proportion that lives in "couple without children", which is increasing and getting later formation.

To summarize this initial section, we have been through individual to household and family dynamics point of view, which Demography has a particular empirical approach. From this base, we can identify changes in demographic components that have being occurring in recent decades, which are well documented in the demographic literature. They have shown that changes in demographic components affects not only the household number in the total population, but also the household size and composition of a population and its subgroups. Additionally, the relationship between mortality, fertility, migration, leaving the parents' house and nuptiality is complex and no "cause and consequence" relationship with household formation and dissolution is easily obtained. The next section will present how does the Family and Household Demography have been trying to develop mathematical and statistical models to understand better the relationship between population and household.

FIGURE 3 – São Paulo's total population by household type and single age, Census 1980, 1991, 2000 and 2010



Source: Author’s elaboration with 1980, 1991, 2000 and 2010 Census data.

1.2 Methods for household and population projections

Before introducing household projection models, it is important to note that the base for a good household projection relies on a good population projection. There are several ways to project population. González and Torres (2012) review the main population projection methods, listed in Table 3. Many of them are based on mathematical extrapolations and others use symptomatic variables, that is, to use variables related to population growth (for example, elector register). Even though those methods are quite useful for plenty of situations demographers daily face, such as projecting total population to calculate rates for a specific period, they have two main disadvantages: they are designed to project only the total population and do not have demographic components as inputs.

On the other hand, the most widespread method for population projection in a country, state or substate level is the Cohort-Component Method, which has been used since the end of eighteenth century (CANNAN, 1895). This is an essentially demographic method, and it depends on a reliable base population and one projection for each component of populational change: mortality, fertility, and migration.

TABLE 3 – List of methods for population projection

Mathematical functions-based methods	Symptomatic variables-based methods
Linear Function	Proportional Distribution
Exponential Function	Vital Rates Method
Logistic Function	Census Ratio Method
Logistic Proportion Estimation	Differential Rate Method
Relative Increment Method	Composition Method
Cohort Relation Method	Ratio Correlation Method
Differential Growth Method	Rates Correlation Method
Apportionment Method (AiBi)	Differential Correlation Method

Source: Based on González and Torres (2012, p. 108).

In contrast to extrapolations methods, the Cohort-Component Method is more dependent on demographic data quality and availability, since there are more sources of errors involved. Bay (2012) lists the main problems for a population projection using the Cohort-Component Method in Latin America. The author states that, since the 1950's, there has been some progress regarding data sources and data quality, but the region still suffers from a lack of regular censuses, the vital statistics systems are still under development (lack of complete coverage) and there are no regular revisions for each component projections, which make a continuous times series impossible to be obtained in some countries.

Another frequent problem is to obtain a correct base population. "*Centro Latino-Americano e Caribenho de Demografia*" (CELADE, 1984) recommends some corrections in raw data to obtain a reliable base population for population projections. The main errors usually found in base populations are: under 5 years old population omission, young adult age groups systematic omission, age heaping, tendency to increase the age among older population, especially after 55 or 60 years. Those distortions can be corrected by ungrouping and several methods that deals with age heaping (MYERS, 1940; SHRYOCK; SIEGEL, 1976; RIZZI; GAMPE; EILERS, 2015). For other population omissions, CELADE (1984) suggests the "*Conciliación Censal*" approach.

Once the base population is obtained, the Cohort-Component Method incorporates all the uncertainties about the trajectory of mortality, fertility and migration. Regarding mortality component, uncertainties relies into how mortality rates will behave at oldest ages (KANNISTO et al., 1994; OEPPEN; VAUPEL, 2002; JANSSEN; KUNST, 2005), the pace of senescence process (WILMOTH; HORIUCHI, 1999; BAUDISCH; VAUPEL, 2012), if there will be a compression process (i.e., a reducing lifespan disparities), or a shifting process of mortality (i.e., shifting the mortality schedule, with lifespan variability remaining nearly constant) (FRIES, 1983; GONZAGA, 2008) and if there is a limit age to lifespan (COLCHERO et al., 2016; OEPPEN; VAUPEL, 2002). Also, there is uncertainties about the gender gap behavior (HORIUCHI, 1999; MOURA et al., 2005; LUY; MINIGUAWA, 2014), the future behavior of external cause mortality at young adults ages and future life gains/loss by causes of death (ABURTO; RIFFE; CANUDAS-ROMO, 2018). Other important issue that is gaining attention since 2020 is the direct and indirect impact of pandemics, such caused by the SARS-COVID 19 virus, on mortality projections and how excess mortality behave according to different countries, ages groups, gender, races/ethnicities, social classes, etc. (GOLDSTEIN; LEE, 2020; LIMA et al., 2020; NEPOMUCENO et al., 2020; DOWD et al., 2020; KASHNITSKY; ABURTO, 2020). Not only in mortality schedules, but many of those direct and indirect SARS-COVID 19 effects may affect fertility and migration components throughout global population.

Uncertainties regarding future fertility rates relies into reducing fertility levels, such as the low-low fertility trap (LUTZ; SKIRBEKK; TESTA, 2006), in societies where gender relations are changing (GOLDSTEIN; SOBOTKA; JASILIONIENE, 2009; MCDONALD, 2000). Also, there is stabilization of adolescent fertility and the increase in fertility at more advanced ages (MYRSKYLÄ; KOHLER; BILLARI, 2011). Migration represents the most uncertain component. It’s difficult to measure and predict. In order to project migration, one should pay attention to return migration and changes in migratory flows. Usually, net migration is estimated by taking into account recent years trend or considering close population where net migration is very low representative compared to natural growth.

In order to project mortality rates in the Cohort-Component Method, Model life Tables (COALE; DEMENY; VAUGHAN, 1983; UNITED NATIONS, 1982) and “Brass logit relational life table model” have been used. In those methods, future mortality rates are estimated by using life Tables with a standard and level patterns that population expect to reach (FRIAS; RODRIGUES, 1981), for example, using recent Latin American Database (LAMBDA) or Human Mortality Database (HMD) projects. However, more recent and sophisticated methods are based on the Lee-Carter family (LEE; CARTER, 1992) and compositional data analysis (BERGERON-BOUCHER et al., 2017), which forecasts mortality rates through observed times series data and the age and time effects, and Bayesian hierarchical models (ALKEMA et al., 2011; FOSDICK; RAFTERY, 2014), which can be found in R package ‘bayesTFR’. Regarding fertility forecast, the oldest models are based on “Brass relational Gompertz fertility model” (BRASS, 1975) and other parameterized models

(COALE; TRUSSELL, 1974; ROGERS, 1986). Recent methods try to adapt Lee-Carter family models to estimate the age and time effects of future fertility rates (EDWARDS et al., 2003).

As we can see, many efforts have been made to improve population projection models and their components. However, the history behind household projection models is more recent than population projection. According to United Nations (1973), the first attempt to project future household number was in 1938 by the States National Resources Planning Committee during the Second World War II to allocate material resources for industrial production. But it was only since the beginning of the 1980s, that Family and Household Demography has grown rapidly and consistently in the international scientific community with the publications of Burch (1979); Bongaarts (1983); Brass (1983); Ryder (1985); Bongaarts; Burch and Watcher (1987); De Vos and Palloni (1989) and Grebenik and Mackensen (1989).

During this period, the International Union for the Scientific Study of Population (IUSSP) held two scientific committees in order to elaborate and define the scope of the Family Demography. The first committee was composed of two international workshops with quantitative focus, looking for models and statistical approaches that resulted in the publications of Bongaarts; Burch and Watcher (1987) and Grebenik and Mackensen (1989). The second IUSSP committee was composed of three meetings with the objective of further exploring the subjective issues of families and households, reported by Prioux (1990); Berquó and Xenos (1992) and Xenos and Kono (1992). As a Family Demography development continuation, one workshop on household formation and dissolution was held in 1984 by the Netherlands Interdisciplinary Demographic Institute (NIDI), which years later resulted in Imhoff et al. (1995) publication.

As discussed in the first part of this Chapter, the main challenge in this field is how to join all the different processes that interfere in the household formation and dissolution in a single model. The processes are complex and interconnected. In that purpose, all these workshops and publications made contributions to discuss and develop models, that tries to “infer hypothetical living arrangements or kin relations from available information on nuptiality, fertility, mortality, and migration and they are tools to disentangle the mechanisms through which the latter affect the former” (DE VOS; PALLONI, 1989, p. 175).

There are several forms to classify and define these models, depending on the analyst perspective. One of them, suggests classifying models into: analytical models, macrosimulation, microsimulation and probabilistic models (WILLEKENS, 2010). The analytical models aim to establish relations between stable population parameters and kin availability, while in the microsimulation models individuals are "units of analysis and transition rates are applied to each person from an initial population, generating individual trajectories and thus a whole distribution" (WAJNMAN, 2012, p. 33). Macrosimulations, on the other hand, "apply transition rates to groups of aggregate individuals, the cohorts, producing trajectories representing average individuals in the cohorts" (WAJNMAN, 2012, p. 33). Finally, more recently probabilistic models try to incorporate variability and estimate confidence intervals for projections.

Another common way to describe a model is from what they can provide us as a result. According to Smith; Tayman and Swanson (2006) and Booth; Hyndman and Tickle (2014), a forecast model distinguishes from a projecting model when the first one results a statistical best estimate of what will happen in the future in terms of likelihood (SMITH; TAYMAN; SWANSON, 2006). Then, a forecast model comprises: a central tendency point (median, mean, …) and an interval forecast, based on variance, including prediction intervals at different levels of probability (BOOTH; HYNDMAN; TICKLE, 2014). On the other hand, projection models are defined as "conditional statements about the future" as "they show what the population would be if particular assumptions were to hold true" (SMITH; TAYMAN; SWANSON, 2006). Because of that, a projection model is formally considered as an exercise, it does not make predictions regarding the future like forecast models (SMITH; TAYMAN; SWANSON, 2006).

Although these definitions are important, this study will follow on the "dynamic" and "static" model definition, which classifies as dynamic, when it studies the cohort behavior over time, or static, when the model studies the future distribution of households based only at a period of time (KEILMAN, 2018; YEPEZ MARTÍNEZ, 2010). Dynamic models usually consider multistate transitions during an individual or family trajectory, so "dynamic family models deal explicitly with family events. A family event is defined as a change in family position that an individual experience during a brief time interval" (KEILMAN, 2018, p. 283). On the other hand, the static models are based on the household distribution (proportions) over a period and statements of those proportions to the future. Usually, the static models are chosen by their modest data demands, but they are also very criticized because the processes of family and household change remain a black box (KEILMAN, 2018). The choose for the dynamic/static definition is due to Brazil and all Latina American countries find themselves in a cross-sectional data context described by Bay (2012), which it's the high dependent data availability is determinant for choosing a projection model. The next section will dig into the previous experiences on household projection, the fields of knowledge that demand this type of information and what have been the choices of these fields about adopting static or dynamic models.

1.3 Fields of knowledge that demand household projections

The usage of population and household projections is placed in a broader context of increasing complexity in governmental and private decisions, where they contribute direct and indirect on public policies formulation, monitoring and evaluation. As Jannuzzi (2012) argues:

> Population projections allow better estimates of the target population to be served by social services in the future. They constitute an important resource to guide the public resources allocation in medium and long-term government plans, define policies nature and content, and establish guidelines for investment in infrastructure and public equipment (JANNUZZI, 2012, p. 89).

In the literature review made in this study and which will be discussed more in the next sections, three major areas stand out as the main fields which demand household projections: housing, family, and environmental studies. In social housing studies, for example, policy makers are interested in the flow of housing deficit and stock measures in order to plan how many housing units will be necessary to supply the social housing demand of a certain region at certain period (GIVISIEZ; OLIVEIRA, 2018; ZENG et al., 2014a). However, the need for new housing units resulting from the demographic dynamics can vary by age and sex. While young adults are exposed to a period of life when formation and dissolution of households is more intense, population aging increasingly places more importance on the elderly households. In Hebei, China, Zeng et al. (2014a) show that compared to 2010, the housing demand of those aged 65 or more will increase by 228.3% and 447.6% in 2030 and 2050. In addition to the number of households by age and sex, the size of these housing units is especially important to optimize the governmental resources allocation in social housing policies (GIVISIEZ; OLIVEIRA, 2018).

To illustrate how important age structure and household dynamic are to housing sector, Martins (2019) details how construction companies and Government housing policies are creating financing programs that consider families life cycle. In these programs, beneficiaries may, at a certain moment, return their apartments in exchange for others with a more suitable plant according to the moment of family expansion and/or retraction. Even so, new buildings follow minimum architectural recommendations, which are reproduced in a standardized manner throughout the country to optimize resources. According to the author, these programs generally target groups of young people and adults up to 40 years which are in stage of family expansion and usually exclude mono-parental households and the elderly households, due to lower income and because the financing time is greater than or equal to the life expectancy of the elderly. The criteria for selecting the target population for social housing policies also focus on specific population that may have different population dynamics and, consequently, different housing needs from total population of that region (APARICIO; BRUSSE, 2018).

Other applications where household projections are considerable relevant is for family studies. In ageing societies, not only the future number of the elderly will matter, but by whom this growing population will be cared for. The elderly care supply is strongly associated with co-residence with spouse and relatives of other generations (ex. children and grandchildren), or with relatives of the same generation (as siblings) who would provide care inside the household. In east Asian countries, such as Fiji, Republic of Korea, Malaysia and Philippines, elderly co-residence with other generations is about 75% and 85% of total elderly care (SOKOLOWSKY, 2008). In developed countries, 1/3 of elderly care is carried out in institutions, while the other 2/3 are carried out within the households (NERI; SOMMERHALDER, 2002). The family is the institution that is responsible for most of the elderly and child care and household projections can give essential information about future kinship availability that will determine the care demand and care supply (CAMARANO; KANSO, 2010).

The future size and composition of households is also important to environmental and consumption studies since the consumption pattern is different between cohorts and it can change over the lifetime of the individuals. Consequently, the participation percentage of certain age groups inside households will have a direct impact on the consumption curve of a population (CASTILLO; PEÑA; GUARDIÁN, 2016; SILVA, 2013). Furthermore, the reduction in household average size has a double impact on the use of resources and on biodiversity (LIU et al., 2003). First, there is an increase in the demand for land and materials (e.g., wood, concrete and steel) needed for housing construction. Secondly, households with fewer residents have less efficiency in per capita resource usage, because goods and services are shared by more people in larger households, such as some fixed environmental costs of energy and water, production and deposition of waste and sewage (LENZEN; MURRAY, 2001; IRONMONGER; AITKEN; ERBAS, 1995; YOUSIF, 1995). Also, some studies verify that the number of cars, consequently the emissions of carbon monoxide and hydrocarbons by cars, are associated with the age of household head and the size of households

(PRSKAWETZ; LEIWEN; O'NEILL, 2004; FIORAVANTE, 2009). In Brazil, more than 4.6 million households were created between 1991 and 2000 due to a reduction in average household size. The contribution of the reduction of the average size in the total number of households is 46% while 54% is due to population growth (LIU et al., 2003). Between 1991 and 2000, the population grew at an annual rate of 1.7% (from 145.6 to 168.3 million inhabitants), while households increased by 3.2% (from 34.7 million to 44.7 million households).

All these research fields demonstrate that household projections must move toward more detail outcomes which take into account not only the total number, but age and sex, size and composition distribution of these future households. However, can the household projections that are being made in Brazil answer to what those fields specifically requires in terms of information? In order to list what can be answered and what blanks exists on the available information, the next section will show a brief worldwide review of the main previous experiences of household projections and a critical review of Brazilian experiences so far.

1.4 Worldwide household projections brief review

The preliminary questions that guided this review focus on: what do these previous projection experiences want to answer? What these household projections will be useful in their field of knowledge? Which methodologies were used to answer that questions? What data sources did they use? What is the population and at what geographic level they applied (municipality, state, country)? Which household classification did they use (if used)? What results did they reach (Total number of households? Disaggregated by sex, age, size, household type classification?).

According to Yepez Martínez (2010), the number and frequency of household projection experiences have been increasing in the last years, however in the 1990’s, only 6 countries had official household projections (Germany, Austria, France, Italy, Netherlands and the United Kingdom) and until 2010 this number had grown to just 10 countries. International institutions responsible for population projections such as the United Nations (UN), United States Census Bureau, World Bank, International Institute for Applied Systems Analysis (IIASA) and the Population Reference Bureau, do not have the practice of household projections.

Yepez Martínez (2010) reviewed the household projections of 68 National Statistical Offices. This effort serves as a starting point for this study. The experiences were compared on method used (dynamic or static), household type classification, data sources, geographic area, and projection year (Table 4 to 8). Tables 9 to 11 are updated information on current household projections practice after 2010's decade including Latin American experiences. Yepez Martínez (2010) found that most of countries still use static methods for projecting households, only Norway, Belgium, Netherland and Japan have had experience with dynamic methods. In general, the models developed after the 1980's Family Demography boom have not yet been absorbed by National Statistical Offices, which still adopt the United Nations recommendation that place the best-known static method (the Headship Rate method), as the most appropriate method for household projection and which had advantages over the other methods available at the time (UNITED NATIONS, 1973). According to Imhoff et al. (1995) the requirement of longitudinal data is the main impediment for dynamic methods to be used since most of the countries still remain in a context of cross-sectional data from censuses and surveys. One consequence of that is the majority countries that have applied dynamic household projection methods are located in European countries where Civil Register allow longitudinal approaches. Also, the investment in paid software and the complexity of these methods which requires a larger set of inputs and calculations are some of the barriers for national and subnational official statistical offices to take a step toward dynamic methods.

Most of the applications occur at national or some federal state level and they have about 10 to 20 years horizon. Regarding household type classification, despite some approximations, they follow the types previously presented here: living alone, couple without children, couple with children, mono-parental and extended household. In the next section, we will see how Brazilian experiences are in relation to these countries.

TABLE 4 – Summary of household projections practice

Country	Institution	Method		Data source	Geographic level	Household classification	Period (horizon)
Czech republic	Czech Statistical Office	Static	Headship Rate	2001 Census	National	i. Couple (married or cohabiting) with children ii. Couple (married or cohabiting) without children iii. Father / Mother only with children iv. Father / Mother only without children v. Multi-personal household without family ties vi- Single-person household	2001-2030
Norway	Statistics Norway	Dynamic	Household Transition	Continuous survey	National	i. Couple (married or cohabiting) with children ii. Couple (married or cohabiting) without children iii. Father / Mother only with children iv. Father / Mother only without children v. Multi-personal household without family ties vi- Single-person household	(NI)
Belgium	National Institute of Statistic	Dynamic	Household Transition	1981 and 1991 Censuses	National	i. Single-person ii. Single-parent iii. Married couple with children iv. Married couple without children v. Unmarried couple with children vii. Unmarried couple without children viii. Others.	(NI) - 2011
France	Institut National de la Statistiquet des Études Économiques (INSEE)	Static	Headship Rate Membership Rate	1990-1999 and 1999-2005 Intercensus data	Regional Metropolitan Department	1. Person living alone 2. People living as a couple 3. Single-parent household 4. Infants 5. Other people (unrelated people living in a household of at least two people) 6. People living in institutional homes.	2005-2030
Germany	Statistisches Bundesamt Deutschland	Static	Membership Rate	2001 Census and Populational data	National and Federal estates	1. Single household 2.Multipersonal: Two people; Three people; Four people; Five people and more	2007-2025

Source: Martinez (2010, p. 221-224).
NI – No Information.

TABLE 5 – Cont. Summary of household projections practice

Country	Institution	Method		Data source	Geographic level	Household classification	Period (horizon)
Australia	Australian Bureau of Statistics	Static	Household Propensity	1986, 1991, 1996 and 2001 Censuses	National and regional	1. Family with both parents 2. Offspring with both parents 3. Single parent 4. Offspring with one parent 5. Couple without offspring 6. Household Family 7. Member of a household group 8. Person alone 9. Person not belonging to a private household	(NI) - 2021
New Zealand	Statistics New Zealand Demography Division	Static	Household Propensity	1986, 1991, 1996 and 2001 Censuses	National and regional	1. Couple without children 2. Person with a couple without children 3. Parent in family with both parents 4. Child in family with both parents 5. Other person in family with both parents 6- Father in single parent family 7. Child in single-parent family 8. Another person in single-parent family 9. Person in multi-personal household 10. Houosehold with a person 11. Person in non-private housing	2001-2021
Italy	Istituto Nazionale di Statistica (ISTAT)	Static	Headship Rate	1961, 1971 and 1981 Censuses	National	(NI)	(NI)
Netherlands	Statistics Netherlands	Mixed model: Static and Dynamic	Membership Rate Household Transition	Census and Civil register	National	1. Living with parents 2. Living alone 3. Living as a couple 4. Single parent / mother 5. Living in an institution 6. Other household member	(NI) - 2050
Austria	National Statistical Institute of Aus tria	Static	Membership Rate	1981, 1991, 2001 Censuses	National	1. Couple without children 2. Couple with children 3. Persons in consensual unions without children 4. Persons in consensual unions with children 5. Single-person households 6. Children of married couples 7. Children of consensual unions 8. Children in single-parent households	2001-2011

Source: Martinez (2010, p. 221-224).
NI – No Information.

TABLE 6 – Cont. Summary of household projections practice

Country	Institution	Method		Data source	Geographic level	Household classification	Period (horizon)
United States of America	U.S. Bureau of the Census - Population Projections Program.	(NI)	(NI)	1990 Census, Current Population Surveys (CPS) 1959-1993.	National and State	1. Married couple 2. Household with female head 3. House with male head 3 – Non-Family Household 4. Single female 5. Single Male	1995-2010
Mexico	Consejo Nacional de Población (CONAPO)	Static	Headship Rate	2000 and 2005 Population counts	National and Federal units	1. Extended 2. Nuclear 3. Other relatives' arrangements 4. Co-residents 5. One-person	2005-2030
Brazil	Centro de Desenvolvimento e Planejamento Regional	Static	Headship Rate	2000 Census	National, estate and municipality	(NI)	2001-2023
Japan	National Institute of Population and Social Securi ty Research	Dynamic and Static	Household Transition Headship Rate	1995 Census	National and regional	1. Single person 2. Couple without children 3. Couple with children 4. Parents with children 5. Other	(NI) - 2020
South Korea	Korea National Statistical Office	Static	Headship Rate	2000 Census	National	1. Married couple with or without child (ren) 2. Father with child (ren) 3. Mother with child (ren) 4. Three generations or more 5. One person 6. Non-familiar 7. Other	2000-2020

Source: Martinez (2010, p. 221-224).
NI – No Information.

TABLE 7 – Cont. Summary of household projections practice

Country	Institution	Method		Data source	Geographic level	Household classification	Period (horizon)
Spain	(1) Instituto Galego de Estatística (IGE) (2) Institut d'Estadística de Catalunya (IDESCAT) (3) Instituto de Estadística de Andalucía (IEA) (4) Instituto de Estadística de la Comunidad de Madrid (IECM)	Static	Headship Rate Headship Rate Household Propension Hybrid Model (1996-2011) Headship Rate (2002-2017)	(1) 1991-2001 Censuses (2) Padrón 1996-2001 (3) Encuesta de Población Activa (EPA)	Autonomous Communities	(1) i. Adult male ii. Adult female iii. Two adults iv. Three or more adults v. An adult with children vi. Two or more adults with children. (3) Couples: i. With children ii. No children; Father or mother only with children; Non-family households; One-person; Pluri-personal (4) i. One-person ii. Multi-person iii. Nucleus only iv. Nucleus + Others	(1) (2002-2017) (2) (2001-2015) (3) (2001-2016) (4) (2002-2017) (1996-2011)
Canada	Statistics Canada	Static	Headship Rate (national) Average household size (local)	1991 Census	Nacional, regional and local	1. Family household: Family with spouse; With children and with/without additional people; Without children and with/without additional people; Household with a father/mother alone: Mother/Father alone with/without additional people 2. Two persons household: Related/Unrelated 3. Three or more persons in the household: Related/Unrelated	1991-2016

Source: Martinez (2010, p. 221-224).
NI – No Information.

TABLE 8 – Cont. Summary of household projections practice

Country	Institution	Method		Data source	Geographic level	Household classification	Period (horizon)
European Union (EU)	Statistics Netherlands, Department of Population.	(NI)	(NI)	Census and Labour Force Survey (LFS)	National: 15 EU countries	1. Living alone 2. Living as a couple 3. Living with one or both parents 4. Other position within the household 5. People without a partner living with children or with other people	1995-2025
England	Office of the Deputy Prime Minister	Static	Headship Rate	1971, 1981, 1991, 2001 Census and Labour Force Survey (LFS)	Regional and Sub-regional	England and Wales use the same classification: 1. Married couple 2. Cohabiting couple 3. Single parent 4. Multiple persons 5. One-person	(NI) - 2016
Wales	National Assembly for Wales	Static	Headship Rate	1971, 1981, 1991, 2001 Census and Labour Force Survey (LFS)	Regional and Sub-regional		
Scotland	Scottish Executive Housing Statistics	Static	Headship Rate	1991 and 2001 Census	Regional and Sub-regional	1. A male person 2. A female person 3. Two adults 4. Three or more adults 5. An adult with one child 6. An adult with two or more children 7. Two or more adults with one or more children	2001-2018
North Ireland	Department for Regional Development. Northern Ireland Housing Executive	Static	Household Propension	Censos 1981, 1991, 2001, Labour Force Survey (LFS) y Encuesta continua de Hogares	National and regional	1. Living alone 2. Households without children 3. Adult only and children 4.Other households with children	2002-2025

Source: Martinez (2010, p. 221-224).
NI – No Information.

TABLE 9 – Summary of the Latin American household projections practice after 2010's

Country	Institution	Method		Data source	Geographic level	Household classification	Period (horizon)
Argentina	Marcos and Módenes (2019)	Static	Headship Rate	1991, 2001 and 2010 Censusse; Population Projection (INDEC)	Buenos Aires Metropolitan Region	Total	2011-2025
Brazil	Giviez (2018) Ministério das cidades	Static	Headship Rate and APC model	2000 and 2010 Census (IBGE)	Municipalities	Total	2010-2040
Cuba	Centro de Estudios de Población y Desarrollo (CEPDE)	Static	Headship Rate	2002 and 2012 Censuses, Population projection (ONEI)	National and provinces	i) One-person household; ii) Basic household; iii) Extended household; iv) Composed by non-relatives	2015-2030
Ecuador	Peter (2015)	Static	Headship Rate, Alpha method, Delta method and Theta method	1990, 2001 and 2010 Censuses. Population projection (INEC)	National and provinces	Total by household size	2010-2020
Mexico	Guardián (2016)	Dynamic	Extended Cohort-Component Method (ECCM)	2010 Censuses, SOMEDE, EDER (2011), ENADID (2009)	National	One generation household: living alone, one person and others, couple only, couple and others Two generation household: couple with children, single parent, and others Three generation households: couple with children and grandparents, Single parents with grandparents, and others	2010-2030
Venezuela	Yépez Martínez (2010)	static	Alpha Method	2001 Censuses, Population projection INE and CELADE	National	Total by household size	2001-2021

Source: Author's elaboration.

TABLE 10 – Summary of the household projections practice after 2010's

Country	Institution	Method		Data source	Geographic level	Household classification	Period (horizon)
China	Zeng et al. (2014)	Dynamic	Extended Cohort Component Method (ECCM)	2000 census and the 2005 mini-census, CLHLS, In-Depth Fertility Survey	National, rural and Urban	One person, Single parent, Cohabiting couple, Married couple	2000-2050
England	Department for Communities and Local Government	Static	Headship Rates	2001 and 2011 using censuses data, ONS sub-national population projections	National	One person, Couple and no other adult, Couple and one or more other adult, Households with dependent children, Other (other multi – person)	2014-2039
England, Scotland, Wales, and Northern Ireland	Welsh Assembly Government WAG (2011)	Static	Headship Rates (Household Membership)	1991 and 2001 Censuses NISRA, DCLG, population projections ONS	National and local government districts.	One person households; One family and no others; A couple and one or more other adults; A lone parent and one or more other adults; Other households	2011-2021
Denmark	Hansen et al. (2013)	Dynamic	Microsimulation model (SMILE)	Danish Civil Registration System and the Housing Register	National and five Danish regions	Owner-occupied housing, cooperative housing, Social housing, public owned rented housing, privately owned rented housing.	1986-2040
Finland and Denmark	Christiansen and Keilman (2013)	Static	Probabilistic household forecasts	Uncertain Population of Europe (UPE), Statistics Denmark and Statistics Finland	National	Married couple, One-person Household, Cohabiting Couple, Lone parent Household, Other private Household, All private households	2007-2037

Source: Author's elaboration.

TABLE 11 – Cont. Summary of the household projections practice after 2010's

Japan	Fukawa (2011)	Dynamic	Microsimulation model (INAHSIM)	Population formed using the INAHSIM model itself	National	One person, Couple only, Couple and children, Single parent and children, Three generation, The others, Newly formed	2010–2050
Korea	Statistics Korea	Static	Headship Rate	2000 Census	National	One-person households, couple-only households and households of couples with children	2017-2047
United States	Zeng et al. (2014)	Dynamic	Extended Cohort Component Method (ECCM)	Census Bureau, NSFH, NSFG, CPS, SIPP, ACS and National Center for Health Statistics	National and by 50 states	One-person, Single parent, Cohabiting couple, Married couple	2000-2050
United States	Feng et al. (2019)	Dynamic	Extended Cohort Component Method (ECCM)	Census Bureau, NSFH, NSFG, CPS, SIPP, ACS, and National Center for Health Statistics	Six Counties of Southern California	One-person, Single parent, Cohabiting couple, Married couple	2010-2040
Spain	Bermúdez Parrado et al. (2011)	Static	Propensity model	Encuesta de Población Activa (EPA)	Andalucía	Couple with children, mono-parental, two or more nucleus, other family households, one person, multi persons	Historical 1990-2012
Spain	Matea (2015)	Static	Headship Rates	2011 Census and Encuesta de Población Activa (EPA)	National	Total	2011-2029

Source: Author's elaboration.

1.5 Critical review of Brazilian household projections experiences

Like most of Latin American countries, Brazil does not have a regular household projection practice. The previous experiences can be summarized in a few academic studies related to the demand for electricity, social housing, housing market and just one projection from an official statistical office. Institutions generally do not publish the methodological part of their studies and when they do, they are difficult to replicate due to lack of information. Therefore, it can be said that this is still a developing area in Brazil.

Comparing the international experiences with the Brazilian experiences in Tables 14 and 15, we can see that unlike some European countries, all experiences in Brazil use static methods. From eleven experiences identified here, six of them use the Headship Rate Method and the other five use the relationship between population and households, called in some studies as "Household Density Extrapolation Method". The data sources used are basically from the last five Census (1970, 1980, 1991, 2000 and 2010) and PNAD survey, reinforcing that we are in a cross-sectional context with few data source alternatives for a longitudinal approach. In some studies, there is a concern regarding methodological improvements for municipality level outcomes, but most of the experiences are still for the national level. In addition, none of the Brazilian experiences has advanced in relation to household type outcomes.

Usually, energy demand studies in Brazil use Household Density Extrapolation Method which is a simple mathematical extrapolation of the relation between population and household number in the last censuses. As a result, most of recent studies is restricted to the total number of households by country region and by household situation (rural or urban). The National Electric Energy Agency (ANAEEL), which published one household projection methodological note in 2004, argues that household projection studies in electricity sector is useful not only to prepare new areas where there will be potential population to be covered by electricity distribution services in the future, but also in areas where service coverage is already universal but still tends to grow. Residential sector is one of the main responsible for the consumption of electricity in Brazil and it is responsible for 23.6% of the electricity consumed in 2011, only behind the industrial sector, with 43.6% (ANDRADE; PINHEIRO, 2014). But no studies give information about why this extrapolation method is chosen or discuss whether this method reaches their objective of effectively plan the energy demand.

Further studies in this field were carried out by Electrobrás (2007), Ministry of Mines and Energy (BRASIL, 2007) and the Directorate of Economic-Energy Studies (DEA) from Energy Research Company (EPE) (BRASIL, 2017). This last household projection study is considered in the official Brazilian electricity consumption projection made by the Ministry of Mines and Energy for 2017 to 2026, according to the National Energy Plan (PNE) (Tables 12 and 13). By the Household Density Extrapolation Method, more expressive household growth is expected in the North and Midwest regions of Brazil, however the weight in relative household distribution will still be much greater in the Southeast and Northeast.

In a different direction from household total number outcomes, some international studies show the impact of age structure and household composition on energy consumption pattern (SÁNCHEZ PEÑA, 2012; IRONMONGER; AITKEN; ERBAS, 1995; O'NEILL; CHEN, 2002; ZAGHENI, 2011; LIDDLE, 2011; CURRAN; DE SHERBININ, 2004). Brazilian studies also draw attention to the importance of considering household composition and age distribution in the future energy demand analysis, since the consumption pattern is different between the cohorts and can change over the life course of the individuals (CRAICE; GUERCIO; D'ANTONA, 2014). Therefore, a simple total number outcome can overlap important subpopulations differences.

TABLE 12 – Projected total population living in households (thousand inhab.) among Brazilian regions, 2017-2026 (Official Brazilian National Energy Plan)

Year	North	Northeast	Southeast	South	Midwest	Brazil
2016	17,822	57,085	86.653	29.542	15.768	206.871
2021	18,885	58,585	89.348	30.483	16.792	214.918
2026	19,799	59,728	91.457	31.232	17.703	219.918
Variation (% per year)						
2016-2021	1.2	0.5	0.6	0.6	1.3	0.7
2021-2026	1.0	0.4	0.5	0.5	1.1	0.5
2016-2026	1.1	0.5	0.5	0.6	1.2	0.6
Participation Structure						
2016	8.6	27.6	41.9	14.3	7.6	100
2021	8.8	27.4	41.7	14.2	7.8	100
2026	9.0	27.2	41.6	14.2	8.0	100

Source: Brasil (2017).

TABLE 13 – Projected total household number (thousand inhab.) among Brazilian regions, 2017-2026 (Official Brazilian National Energy Plan)

Year	North	Northeast	Southeast	South	Midwest	Brazil
2016	4,763	17,042	28,999	10,389	5,242	66,435
2021	5,296	18,282	31,380	11,371	5,890	72,219
2026	5,819	19,449	33,662	12,333	6,536	77,799
Variation (% per year)						
2016-2021	2.1	1.4	1.6	1.8	2.4	1.7
2021-2026	1.9	1.2	1.4	1.6	2.1	1.5
2016-2026	2.0	1.3	1.5	1.7	2.2	1.6
Participation Structure						
2016	7.2	25.7	43.6	15.6	7.9	100.0
2021	7.3	25.3	43.5	15.7	8.2	100.0
2026	7.5	25.0	43.3	15.9	8.4	100.0

Source: Brasil (2017).

In housing studies, the attention is usually focus on concepts of housing deficit and household demographic demand. Housing deficit is related to the stock of households while household demographic demand is related to the flow (demand) of housing needs over the time horizon. If the supply of new housing exceeds the housing demand (flow) in the projection horizon, there will be a tendency to reduce the deficit (stock) (ABRAINC, 2018).

The Headship Rate Method is often used to calculate household demographic demand. Givisiez; Sawyer and Rios-Neto (2006) used the Age Period Cohort (IPC) model to estimate future trends in headship rates. In a second study, Givisiez and Oliveira (2014) use regression models to estimate headship rates for small population in Brazil, in which the observed headship rate does not have a continuous behavior as observed in aggregated data from countries, states or large regions. Also, in this field, the Brazilian Association of Housing Market Developers (ABRAINC) and Getúlio Vargas Foundation (FGV) use Household Density Extrapolation Method to project the Brazilian housing market between 2018 and 2027, the number of families was estimated according to different income groups. Compared to energy demand studies, Brazilian housing studies are more concern about household distribution according to sex, age and small areas.

As the only experience of an official statistical office, SEADE Foundation (part of the São Paulo State Government Secretariat) published the “Population Projection System” in 2017, an important contribution to population and household projections in Brazil. According to the methodological notes, Population Projection System is a Cohort-Component Method application combined with the Headship Rate Method for the total population and household projection in São Paulo State. The population projection is disaggregated according to five-year age group, sex, school age groups and urban and rural situation for the 645 São Paulo municipalities, 96 capital districts and the administrative regions of the State, in the period from 2011 to 2050. The data sources used are from the Civil Register Offices of all municipalities in São Paulo State, Census and Vital Statistics Register. Meanwhile, the household projection obtained by the Headship Rate Method is only performed for the total number of occupied households in each municipality, metropolitan regions, administrative Regions, Health Regional Departments or São Paulo State as a whole. There is no information about age, sex, or household type.

Despite some studies, Brazilian experiences are in general a step behind from international experiences in detailing the outcome information by geographic level, household type, age, and sex distribution. Additionally, in Brazilian static methods application, the time series are generally shorter (using few points in time) such as the last four Census or population counts, which flaws a clear view of headship rate or household density change pattern over time. None of the studies have deeply discussed the effect of changes in headship rates by sex, age or household type and where those trends are going to.

All these research areas demonstrate that household projections must move toward more detail outcomes which take into account not only the total number, but the age, sex, size and composition distribution of these future households. But the question is, which demographic, mathematical or statistical model is able to express the complex relation between population and households and produce such detailed information of the future?

TABLE 14 – Summary of the current Brazilian household projection practice

Year	Institution	Method		Data source	Geographic level	Household classification	Period (horizon)
2004	Empresa de Pesquisa Energética and ANAEEL	Static	Headship Rate	2000 Census, 1996 Population Count and 2004 IBGE's population estimative	Municipality (rural and urban)	Total (Rural and Urban)	(2000-2004)
2005	Leon and Pessanha	Static	Population/Household Ratio trend extrapolation	1980, 1991 and 2000 Censuses	National	Total	
2006	Givisiez, Sawyer and Rios-Neto	Static	Headship Rate and APC modeling	1970, 1980, 1991 and 2000 Censuses	Belo Horizonte Metropolitan Region	Total by sex	(2000-2010)
2007	Eletrobrás	Static	Headship Rate and APC modeling	1970, 1980, 1991 and 2000 Censuses	National	Total	(2000-2020)
2007	Ministério de Minas e Energia	Static	Population/Household Ratio trend extrapolation	1970, 1980, 1991 and 2000 Censuses	National	Total	(2000-2030)
2009	Convênio Cedeplar, Secretaria de Habitação do Ministério das Cidades	Static	Headship Rate and APC modeling, average household size	1970, 1980, 1991 and 2000 Censuses	National, States and Municipality	Total / Total by sex	(2010-2023)

Source: Author's elaboration.
NI – No Information.

TABLE 15 – Cont. Summary of the current Brazilian household projection practice

State	Institution	Method		Data source	Geographic level	Household classification	Period (horizon)
2014	Givisiez and Oliveira	Static	Headship Rate and Polynomial regression	2000 Census and Sistema Nacional de Nascidos Vivos (SINASC)	Municipalities from Acre and Rio de Janeiro	Total by sex and age	(NI)
2014	Andrade and Pinheiro	Static	Population/Household Ratio trend extrapolation	PNAD 2009	National	Total	(2005-2020)
2017	Fundação SEADE	Static	Headship Rate	Civil Register, SINASC, SIM and Census	São Paulo's state municipalities	Total	(2017-2030)
2017	Estudos Econômicos-Energéticos (DEA)	Static	Population/Household Ratio trend extrapolation	IBGE	National	Total	(2017-2026)
2018	Associação Brasileira de Incorporadoras Imobiliárias (ABRAINC) and Fundação Getúlio Vargas	Static	Population/Household Ratio trend extrapolation	PNAD and IBGE projections	National	Total	(2018-2027)

Source: Author's elaboration.
NI – No Information.

CHAPTER 2 – A COMPARISON BETWEEN THE EXTENDED COHORT-COMPONENT METHOD AND THE HEADSHIP RATE METHOD

2.1 Combining the cohort-component method and the headship rates method: the classical approach

The classical static approach comprises combining the Cohort-Component Method for population projection by age and sex with the Headship Rate Method for household projection. The Cohort-Component Method has been used for over 100 years by demographers and in related areas. The first use of the method was by Cannan (1895), but its formal formulation is attributed to Whelpton in 1936. According to UNFPA, "[…] the procedure involves surviving each age cohort, adjusting for migration, calculating the number of children born in the projection interval, and adjusting for their mortality and migration" (UNFPA, 2015, p. 20) and "normally carried out using a software developed specifically for this purpose" (UNFPA, 2015, p. 20). Following the logic of the demographic balancing equation, the future population of a cohort aged x years old at time $t + n$ will be defined as the current population of that cohort at time t, adding the births $B(t)$ and immigrants $I(t)$, and subtracting deaths $D(t)$ and emigrants $E(t)$ of that cohort:

$$P(t + n) = P(t) + B(t) - D(t) + I(t) - E(t) \quad (1)$$

Therefore, the data required for a population projection using the Component Cohort Method comprises a base year population by age and sex, sex-specific life Tables, age-specific fertility rates and age-sex-specific net migration for each interval in the projection period (unless one is assuming that the population is closed to migration) (UNFPA, 2015). The great challenge of this method and what will determine the quality of the projection is how to project the age-specific mortality, fertility and migration for each period. A whole literature has been produced in each area of demography to answer that question and many models have being developed.

Once having projected the population by the Cohort-Component Method, the Headship Rates Method (HRM) is defined as the multiplication between the headship rates calculated for each household type of the population and the projected population in the period. The headship rates $h(i, x, t, d)$ of sex i, age x, in the period t and household type d, are defined as the ratio between the total households heads $H(i, x, t, d)$ of sex i, age x, in the period t and household type d, by the population exposed to the risk of being household head $P(i, x, t, d)$ of sex i, age x, in the period t and household type d, (KONO, 1987):

$$h(i,x,t,d) = \frac{H(i,x,t,d)}{P(i,x,t,d)} \quad (2)$$

Applying the approach of the HRM used in this work, we have that the future number of household heads in k years ahead $H(i,x,t + k)$ by sex i, age x, in period $t + k$ for the household type d, is given by:

$$\sum_d \sum_j \sum_i H(i,x,t + k,d) = \sum_d \sum_j \sum_i P(i,x,t + k) * h(i,x,t + k) \quad (3)$$

Where $P(i,x,t + k)$ is the population of sex i, age x, in period $t + k$ and household type d, projected by the Cohort-Component Method; and $h(i,x,t + k)$ are the headship rates estimated by sex i, age x, for period $t + k$ and household type d.

2.2 Issues concerning the Headship Rate Method

Even though the Headship Rate Method requires a simple set of calculations and estimates which are available in most of the cross-sectional census and surveys in Brazil with good data quality, this approach should be applied with some caution. Many authors have criticized this method over time (BURCH, 1999; MASON; RACELIS, 1992; MURPHY, 1991; ZENG et al., 2014b). For Burch (1999), the major problem with the Headship Rates Method is to consider all persons who are not households head in a single “non-head” group with little or no information about them. As we multiply just the proportion of those who are head to the projected population, the other members of household are just indirectly projected.

Another criticism of the Headship Rates Method is that there is no theoretical link between headship rates and demographic rates. Because of that, it is not possible to incorporate into household projection assumptions of demographic processes related to the growth trend or decrease of headship rates (LEIWEN; O 'NEILL, 2004; MASON; RACERIS, 1992). For example, it is not easy to determine what will happen in terms of the headship rate of a given population if the fertility drops from one period to another. In addition, the “static nature” of the Headship Rate Method, which gathers information only about the period makes it impossible to reveal differences among cohorts over the time (ZENG et al., 2014b).

Also criticized are the censuses and surveys methodology that identifies the household head. In Brazil, the Portuguese term used to identify the household head (*chefe*) was vague and could be interpreted as referring to the man economically responsible for the household (known in the literature as “breadwinner”) (SABÓIA; SOARES, 2012; VIEIRA, 2017; CAVENAGHI; ALVES, 2018). It was just in 1980 that it became clearly apparent in the census instructions that the household head could be a woman (SABÓIA; SOARES, 2012;

VIEIRA, 2017; CAVENAGHI; ALVES, 2018). Since that census there was also a differentiation between the head of household and the head of the family to distinguish households composed by more than one family. Since 2000's census, the term has changed to "responsible for the household", also a vague term but without the gender connotation of the former. And 2010's census included an option to "shared responsible for the household" (VIEIRA, 2017).

For some authors, all these methodology changes lead to a misunderstanding so that it is not possible to know if an increase of certain household type between two periods occurs due to a real increase of that specific household type or due to a cultural change in the census question understanding (ZENG et al., 2014b; MURPHY, 1991). Consequently, calculating headship rates based only on these census questions could carry inaccuracies. However, using methodology changes from 1991 to 2000 in Brazilian census, Sabóia and Soares (2012) compared to PNAD results although the headship definitions are different, the results are quite consistent (Figure 44, Appendix).

2.3 The Extended Cohort-Component Method (ECCM)

Based on increment-decrement life Tables, the Extended Cohort-Component Method (ECCM) proposed by Zeng; Vaupel and Wang (1997; 1998) and Zeng et al. (2006) can model the future number, size and composition of a household using transition rates between demographic states which are related to household formation and dissolution. It includes age, sex, marital status, parity, number of co-resident children, number of co-resident parents, type of household (private or institutional) and rural/urban, race or ethnicity (optional).

In this method, the population composition is modified by a multiplication between the transition rates and the corresponding age and sex surviving population at each period. As the composition of the population changes over the projection time, the number, size and composition of households also modifies, giving the method a dynamic aspect. In contrast to the static methods where the headship rates are obtained independent from the population projection and based in period information, all the household information in ECCM is obtained intrinsically from the cohort population dynamics. The data required comprises a census micro data file at initial year as a base population, age-specific demographic standard schedules – including the transition rates – and the projected, or assumed, demographic summary parameters for the future years (Scenarios).

The basic idea of the method can be summarized in 4 steps. First, define a base population at time t. Second, update the demographic states of the base population (from time $t, t + 1$). Third, summing up the households using the new demographic states ($t + 1$) and finally, return to the first step, where the new updated population ($t + 1$) will be the new initial population (t).

The update equations (step 2) follow the principle by which the population in a given demographic state depends on the population in the same state in the previous time, adding the population who entered into the demographic state, minus the population who left the demographic state:

$$N_i(x + 1, t + 1) = N_i(x, t) + E_i(x + 1, t + 1) - S_i(x + 1, t + 1) \qquad (4)$$

Where $N_i(x + 1, t + 1)$ is the number of people aged $x + 1$ that are in the state i at time $t + 1$; $N_i(x, t)$ is the number of people aged x at time t into state i; $E_i(x + 1, t + 1)$ is the number of people aged $x + 1$, at time $t + 1$, that entered at state i during the period $(t, t + 1)$ and $S_i(x + 1, t + 1)$ is the number of people aged $x + 1$, at time $t + 1$, that left state i, during the period t and $t + 1$. The population that entered and the population that left each demographic state is obtained by the life-table principle of multiplying a probability transition by the surviving population:

$$l_j(x + 1, t + 1) = \sum_{i=1}^{T} l_j(x, t)\boldsymbol{P_{ij}}(x, t) \qquad (5)$$

Where $l_j(x + 1, t + 1)$ is the number of people aged $x + 1$ at time $t + 1$ which have transited from state i to the state j; $l_j(x, t)$ is the number of people aged x at time t who are in state i ($i = 1,2,3 \dots . T$) and $\boldsymbol{P_{ij}}(x, t)$ is the probability matrix that a person of age x at time t who is in state i transit to state j at age $x + 1$ and time $t + 1$. However, as Zeng et al. (2014) argue, it is computationally difficult to estimate the probability matrix $\boldsymbol{P_{ij}}(x, t)$ considering all the variables of the method: seven marital states/union, three states of co-residence with parents, six parity states and six cohabiting states with children. Consequently, the dimension of the probability matrix $\boldsymbol{P_{ij}}(x, t)$ is very large. To solve that, the authors propose a computational strategy to calculate the matrix which prevents the problems of estimating very high dimensional Tables of cross-status transition probabilities and can be found in Zeng et al. (2014b).

The "accounting equations" (step 3) that summarize all the household information from the updated population combine the reference person with the co-residence with parents, the co-residence with children, parity status, sex and marital status in order to compose the households. The accounting equations are shown in Zeng et al. (2014b) and can be illustrated in their words:

> For example, a married or cohabiting woman who is not co-residing with parents and whose number of co-residing children is 0 ($c = 0$) is a reference person representing a one-couple only household; if her number of co-residing children is greater than 0 ($c > 0$), she represents a two-generation & couple household of $2 + c$ family members ($c > 0$). If the reference person is not married and not cohabiting (can be a man or woman), he or she represents a single-person alone ($c = 0$) or single-parent household of $1 + c$ family members ($c > 0$) (ZENG et al., 2014b, p. 22).

All the steps of the method, including update and accounting equations, procedures for ensuring consistencies between the two sexes and between parents and children are implemented in the software called ProFamy and are explained in detail in Zeng et al (2014b).

Moreover, Zeng et al. (2014b) compared the 50 U.S. states and Washington DC household projection between the ECCM and the HRM. The comparisons between the 2000 census data and the housing units' projections show that the HRM forecast error is substantially larger than the ECCM. However, even though the ECCM is a consistent theoretical advance and more realistic model compared to the HRM, the data quality and the data sources availability in U.S display a different reality from the Brazilian restricted cross section data context. In the next sections, we will investigate how the data quality and data source availability can affect the projection according to the different natures of the method.

2.4 Issues concerning estimating age-status-specific demographic rates for ECCM in Brazil

The ECCM requires a considerable amount of age-status-specific demographic rates estimation in order to project households. These transitions rates are calculated by dividing the number of occurrences by the exposure risk as defined below:

> Transition rates relate the number of events during an observation period to the duration at risk during that period. The duration at risk is measured in person days, person-months, or person-years. Transition rates, as defined here, are also known as occurrence-exposure rates because they relate the number of occurrences in a period to the exposure time in that period (WILLEKENS; PUTTER, 2014, p. 384).

However, this information is not always easily available and requires some procedures to correct poor data quality of event counts, especially by single age, parity, and marital status. Here, will be shown the most problematic concerns about calculating these transition rates using Brazilian data sources.

First of all, Profamy's software requires a base population to which the transition rates will be applied. The user needs to prepare an ASCII (.DAT) format file as a microdata database. Information about age, sex, relation to the head of household, household identifying number, marital status, parity, race, rural/urban are needed following the ProFamy's category definitions. However, the categories defined by the software do not always correspond to the categories found in the census and some errors can be caused by the matching recodification. For example, in the 2000 Brazilian Census, the question about "relationship with the household head" had a set of 11 categories to classify the residents of the household and in the 2010 Census, the same question had 20 categories, while the Profamy requires only 8 categories. In the 2000 Census, there is no category "grandfather" or "great-grandfather". So, those persons must be included in the category "other relatives" in Profamy's categorization, creating future problems in estimating 3 generation households.

When estimating mortality standard schedules, there are three data sources available in Brazil, the Demographic Census, the Civil Register, and the Mortality Information System (SIM). Mortality estimation in remote regions is still a challenge in Brazil, resulting in incomplete death registers. Consequently, both the Headship Rate Method and the Extended Cohort-Component Method depend on indirect mortality estimations in areas where the coverage and data quality of death registers differs by sex and age. Methods such as Growth Balance (BRASS, 1975), The Synthetic Extinct Generations Method (BENNETT; HORIUCHI, 1981) and Adjusted Extinct Generations (HILL; CHOI, 2004) must be applied in order to correct the death under-estimation. However, all these methods have assumptions, such as a stable and closed population, that are easily violated and may result in errors mainly to estimate infant and elderly mortality.

Besides these difficulties, estimating mortality by Census may present several problems of death under-enumeration, depending on the respondents' statements, such as: the possibility of household dissolution after a member's death, considering persons as members of more than one household at the same time, errors in the declaration of the reference date and errors regarding age declaration of both the living and the dead. In addition, the census question about “deaths occurring at home in the last 12 months”, which allows direct mortality estimation for all ages, was only present in the 1980 and 2010 Censuses, not allowing the use of this data source for different projection base population years.

Similar to mortality, the fertility estimation can be derived from the Demographic Census, the Civil Register or the Live Births Information System (SINASC). The Extended Cohort-Component Method requires not only the total fertility rate, but the total fertility rates by parity, the transition rates between parity states by marital status, mean age at first child and sex ratio at birth. So, not only good fertility data quality is required, but also good joint information about marital status and parity. Neither Civil Register nor SINASC include in their databases consensual union as one of the marital status, consequently there is a mismatch with other marital status-specific demographic rates that consider the consensual union in the Extended Cohort-Component Method. Therefore, Census is the only data source which gathers all the information needed.

However, estimating fertility by Census also has problems. First, women tend to omit information about their children, especially those who live in different households or who already have died. Second, errors can also occur due to the classification of children born dead as a live birth. “Reference period” errors also affect the data and occur when the respondent omits part of their fertility because they consider that the birth occurred in a different period from the census question. And finally, mistakes may occur because the interviewers classify the total number of children equal to zero by not completing the information, considering that contingent of women as missing values. To solve those problems, indirect fertility estimation by P/F Brass Method (BRASS, 1974) is required. But in the same way as mortality indirect estimation, this indirect method also has assumptions, such as a stable population and a context of non-declining fertility that are easily violated and may result in errors.

More complex than mortality and fertility, the marital status transition rates estimation is the main obstacle to the Extended Cohort-Component Method application. In Brazil, few studies were made about transition between marital states, all of them just at the regional level and not considering consensual unions (CORTEZ, 2007; AUGUSTO, 2004; FREIRE; AGUIRRE, 2000; FREIRE et al., 2006).

Cortez; Lazo and Magalhães (2008) argue that, since the National Household Sample Survey (PNAD) eliminated more detailed questions related to the marital status in

1996 and the Demographic Census dropped the retrospective question related to marital status after 1991, the Civil Register became the only national data source for studying marital status transitions. Nevertheless, it is not possible to directly apply the methodology of the multistate life table to estimate the transition rates between marital status: in the Civil Register there is no record of widowhood, that is, there is no continuous record that counts the transition from the married to widowed state, so the number of widowed persons is obtained through an indirect measure using information on mortality. In addition, civil registration concerns only legal unions. Although legal unions still represent most of all unions, historically in Brazil and Latin America, there is a high proportion of consensual unions. According to Quilodrán (2008), in Brazil people in consensual unions increased from 7.1% to 34.5% between 1970 and 2000. This growth of four times the number of people in consensual unions in 30 years represents the largest growth of unions in Latin America. The Extended Model of Cohort-Components would have its potential harnessed much better if the option of 7 marital states were used (marital status including consensual unions). Not considering them is a serious mistake in a Latin America context.

However, as civil registration concerns only legal unions, the only way to include the consensual unions in the model was estimating the transition rates between 4 legal marital states (Single, Married, Divorced and Widow), but considering as “marriage” also the individuals who are in a consensual union. In the same way, “single” is considered as both individuals never married and those who were not in a consensual union.

Here, we are assuming that people who are in a consensual union have the same transition rates and age pattern behavior of people in legal union, but that is not true. Again, these approximations yield errors that can propagate in each year of projection.

Furthermore, most of the age-status-specific demographic rates estimation for the Extended Cohort-Component Method uses single year age and sex distributions. Such detailed information requires procedures for smoothing or disaggregating data and also can add errors to the projection. For example, to calculate single year age transition rates between parity states and by marital status, we end up with this case: how many 15-yearold widowed women had transited from 3 to 4 children state during the last year in the São Paulo? The answer is: very few. Consequently, if during that specific year there were randomly more cases than usual in previous years, the smoothing data procedure will produce a totally different single age and sex distribution, overestimating the number of widowed women for each year of the projection.

In conclusion, the Extended Cohort-Component Method has a significantly increased number of variables compared to the Headship Rate Method that, on one side, gives more realistic complexity to the model, but on the other side, adds a larger set of error sources due to different data sources' quality. The difference in the number, size, and composition of projected households compared to the real census tabulation carries the sum of all type of estimation errors: coverage errors and poor data quality of data sources, violation of assumptions of methods for data correction, smoothing, approximations and data aggregation. Thus, improving the Extended Cohort-Component Method application in Brazil is related to improving the data availability and data quality in Brazil.

The general solution suggested by Zeng et al. (2014b) to overcome all these distortions is to borrow data from other countries or regions which have better data quality. For example, to solve the lack of widowhood data in Brazil, we could use the transition rates from the United States. But some natural questions that come up from these substitutions are, does any place in the world have a marital status pattern similar to Brazil, when even Latin America has different patterns? How far from reality we are each time we replace a transition rate? What are the impacts of this substitution in the projection? Here, a household projection will be done to evaluate whether Brazil has enough data source availability and data quality to support a step forward to a dynamic approach compared to the classic static approach.

2.5 Material and methods: input data required and data sources

An efficient and well-known way of evaluating a population or household projection method is to perform a projection between two past periods for which observations are known and compare the projected results to observed census data. Previous studies consider acceptable errors for projections at the national level at a magnitude of approximately 2-5% errors for the total population and 5-10% prediction errors for subpopulations over a 10-year period (ZENG et al., 2014b; CAMPBELL, 2002; ESRI, 2007; KHAN; LUTZ, 2008).

For the Headship Rate Method application, the software POPGROUP©[1] was used. Developed by Bradford Council, the University of Manchester, and Andelin Associates, it is based on Microsoft Excel through a series of VBA macros that form a set of demographic models capable of projecting population and households by specific geographic areas and

[1] POPGROUP 4.1 © software license was provided by Ludi Simpson in 02/2017, Edge Analytics Ltd. Leeds Innovation Centre | 103 Clarendon Road | Leeds | LS2 9DF 0113 384 6087/6088 | www.edgeanalytics.co.uk.

population groups. This software is used officially by the UK National Statistical Agency for local governments in England, Scotland, Wales, and Northern Ireland (WAG, 2011).

For the Extended Cohort-Component Method application, the software ProFamy©[2] developed by Zeng; Vaupel and Wang (1997; 1998) and Zeng et al. (2006; 2014b) was used. It is an improvement of multi-state life tables software called FAMY (ZENG et al., 1990), based on its predecessor FAMTAB, developed by Bongaarts; Burch and Wachter (1987) and Willekens and Putter (2014).

To evaluate whether Brazil has enough data source availability and data quality to support a step forward a dynamic approach, two household projections were carried out using the same projection horizon and same scenarios schedules. One historical projection from 2000 to 2010 compared ECCM and HRM with 2010 Census observed data, while a 2010 to 2050 projection compared the behavior of each method into the future.

São Paulo state was chosen as base population because of its satisfactory data coverage as Brazil has great differences in terms of data availability and data quality by federal units. Some of the results from the methods explored in this work, mainly when exploring projection errors, could face problems due to lack of coverage and underreporting deaths and live births records. From 2000 to 2010, deaths and births registration in São Paulo state reached a level equal to or greater than 90% coverage (PAES, 2005; CGIAE, 2011). Despite of that, São Paulo State is one of the greatest economic development areas of the Latin American's largest country. With an advanced stage of the demographic transition process, its population grew from 25 million to 41 million inhabitants in the last four decades, corresponding for about 20% of the Brazilian total population according to Census. As a result of this dynamic, the Aging Index triplicates, from about 20 to almost 65 people aged 60 or over for every 100 people aged 0 to 14-year-old (AIDAR et al., 2017; SEADE Foundation, 2017; BERQUÓ et al., 2014). Such rapid demographic transformations have been changing the size, composition, and organization of the 12,837,281 São Paulo's households (AIDAR et al., 2017).

Input data required and data sources

The data used by each household projection method is presented in Table 16. Although the methods require different inputs information, the mortality, fertility, and

---

[2] PROFAMY © software license was provided by Yi Zeng and Zhenglian Wang in 08/2016 | 16 Suzhou Street, Digital China Building, 7F | Copyright © 2018 ProFamy 京ICP备17060793号-6| www.profamy.com.cn

migration estimate from 2000 to 2010 for São Paulo state are the same, ensuring a common starting point for both projections.

TABLE 16 – Inputs and data sources required for each household projection method

Input information	Cohort Component + Headship Rate (PopGroup)	Extended Cohort-Component (ProFamy)
Base Population	Simple age and sex; Source: microdata Census Microdata 2000 and 2010 Population estimates from 2001 to 2009 by five-year age groups and sex; Source: IMP - SEADE Foundation	Microdata containing age, sex, relation to the head of household, household identifying number, marital status, parity; Source: Census Microdata 2000 and 2010
Fertility	Number of live births for each year (2000 to 2010) Source: SINASC / DATASUS Specific fertility rate by five-year age and sex; Source: SINASC / DATASUS	Total fertility rate; Source: SINASC / DATASUS Total fertility rate by parity; Census Microdata 2000 and 2010 Transition rates between parity states by marital status; Census Microdata 2000 and 2010 Mean age to first child; Census Microdata 2000 and 2010 Sex Ratio at birth; Census Microdata 2000 and 2010
Mortality	Number of deaths 5 years age-group and sex; Source: SIM / DATASUS Specific mortality rate by single age and sex; Source: SIM / DATASUS	Probability of surviving by simple age and sex; Source: SIM / DATASUS Life expectancy at birth; Source: SIM / DATASUS
Migration	Percentage distribution of immigrants and emigrants by five-year age group. Source: Census Microdata 2010.	Annual net migration by sex and simple age; Source: Census Microdata 2010
Marital status	-	Transition rates between marital states by sex and simple age; Source: Civil Register 2000 (Cortez, 2008) Mean age at first marriage / union; Source: Census Microdata 2010 General rate of marriages and divorce; Source: Civil Register 2010 Absolute divorce number; Source: Civil Register 2010
Household	Headship rates by household type for 2000 and 2010; Source: Microdata Census 2000 and 2010 Population living in collective or non-private households; Source: Microdata Census 2000 and 2010	Average age when leaving parental home by sex and age; Source: Census Microdata 2010 Proportion of persons aged 45 to 49 years who do not live with their parents by sex; Source: Census Microdata 2010 Proportion of elderly living with children by sex; Source: Census Microdata 2010 Average number of other relatives and non-relatives by sex; Source: Census Microdata 2010

Source: Author's elaboration.

2.6 Trends, assumptions, and scenarios

As mentioned by Givisiez and Oliveira (2018), the key to a good population projection is adequate hypothesis formulation about the future behavior of the demographic components. The assumptions about demographic components requires some knowledge about the past and present behavior to design a plausible and coherent trend to be achieved in the future. To ensure adequate hypothesis formulation, this study uses as a baseline recent population projection from SEADE Foundation (Official São Paulo State Statistic Office) and IBGE (Brazilian National Statistic Office). Only mortality, fertility and migration assumptions will be considered, all the other variables related to household formation and dissolution will be considered constant from the base year of the projection.

The Total fertility Rate (TFR) in Brazil experienced a rapid drop from around 6 children per woman in 1960 to 1.9 in 2010. Oliveira and Fernandes (1996) showed that the historical TFR annual trend can be modeled by a logistic curve with a lower limit equal to 1.7, which has been confirmed by more recent data (Figure 45, Appendix). São Paulo state, in contrast, has historically lower TFR levels compared to the total population, the reason why SEADE Foundation believes that São Paulo TFR can reach 1.5 children per woman in a long-term horizon:

> […] the fertility of women in São Paulo is already very low, lower than the level of population replacement since 2000. The average number of children per woman in the state was 1.68 children in 2016 and the hypothesis formulated was reduction of this indicator in the future, reaching 1.5 children in 2050 (WALDVOGEL; CAPASSI; MORAIS, 2018, p. 8).

IBGE's scenario is more conservative and projects São Paulo's TFR of 1.65 children per woman in 2050. The maximum TFR scenario will be based on fertility recovery in countries that have ever had a TFR below 1.3 or 1.5. It follows the example of New Zealand, the United States and France, that have had a lower TFR between 1.6 and 1.8 (the same level as São Paulo state today) and recovered to replacement level (GOLDSTEIN; SOBOTKA; JASILIONIENE, 2009).

In the latest review of mortality trends in Brazil, IBGE argues that the life expectancy at birth in Brazil increased from 45,5 years (42,9 – men; 48,3 – women) in 1940 to 73.9 years (72,5 – men; 79,6 – women) in 2010. While São Paulo state historically has a higher life expectancy at birth compared to the total population (IBGE, 2018). SEADE's scenario believes in a slow increase in life expectancy to a maximum limit of 81.7 years:

> In 2016, the life expectancy of São Paulo population was 75.8 years and it is expected that the projection horizon will be 81.7 years in 2050. Mortality may also vary depending on the composition and/or evolution of the different causes of death that affect São Paulo population. However, the most likely behavior is the increase of average life span, even if it may be slow (WALDVOGEL et al., 2018, p. 6).

The scenarios created here will follow SEADE's trend that considers a decrease in the gender gap in life expectancy, which means a gender gap of 7 years observed in 2010 for São Paulo state will drop to 5.13 years in 2050 (IBGE, 2018; WALDVOGEL; CAPASSI; MORAIS, 2018). In contrast, the scenario for mortality worsening will consider the unlikely decline in life expectancy at birth over time, which supposes that new cohorts will face a worse lifestyle, higher prevalence of obesity, diabetes, cancer, and cardiovascular diseases, increase of violence, higher pollution levels and/or a return of some infectious disease pandemics (OLSHANSKY et al., 2005). Only 1 year of reduction in life expectancy will be considered.

Finally, the migration component which had an important role in São Paulo state during the 70's and 80's, nowadays its volume has been declining and losing considerable weight in the total growth rate. The annual net migration, corresponding to the period 2000 to 2010, was the lowest ever recorded in recent history: 47,265 people and 1.21 migrants per thousand inhabitants, respectively (WALDVOGEL; CAPASSI; MORAIS, 2018). Therefore, São Paulo population will be consider closed to migration.

> It is also believed that the net migration won't register high values and relative weights in the demographic dynamics of São Paulo as it was in the past. (...) it is possible to mention, for example, the drop in fertility recorded throughout the country, which has meant that the migration of large families is no longer expected today; the existence of new attraction poles in the Brazilian territory, which represent other options for people who think or need to leave their regions; the high costs of living and housing market that may negatively impact the population attraction for certain regions in the state (WALDVOGEL; CAPASSI; MORAIS, 2018, p. 9).

Five scenarios were created combining high/low mortality, high/low fertility and constant rates and they are presented in table 17 below. The first scenario considers that life expectancy reaches the expected maximum, but the TFR has a recovery to the replacement level. The second scenario is the most likely scenario, where the TFR continues to drop and reaches the expected minimum, while life expectancy at birth reaches the expected maximum. Scenario number three is the trend scenario that shows what will happen if the base year rates stay the same until the end of the projection. And finally, the fourth and fifth scenarios project a drop in life expectancy at birth, contrasting the scenario of low fertility and high fertility levels.

TABLE 17 – Five horizontal scenarios made for 2050, São Paulo, Brazil

2050 Scenario	Total Fertility Rate (TFR)	Life Expectancy at birth (e0)
1. Maximum e0 + Maximum TFR	2.1	84,2 – Women 79,07 – Men
2. Maximum e0 + Minimum TFR	1.5	84,2 – Women 79,07 – Men
3. 2010 Constant rates	1.7	79,3 – Women 72,3 – Men
4. Minimum e0 + Maximum TFR	2.1	78,3 – Women 71,3 – Men
5. Minimum e0 + Minimum TFR	1.5	78,3 – Women 71,3 – Men

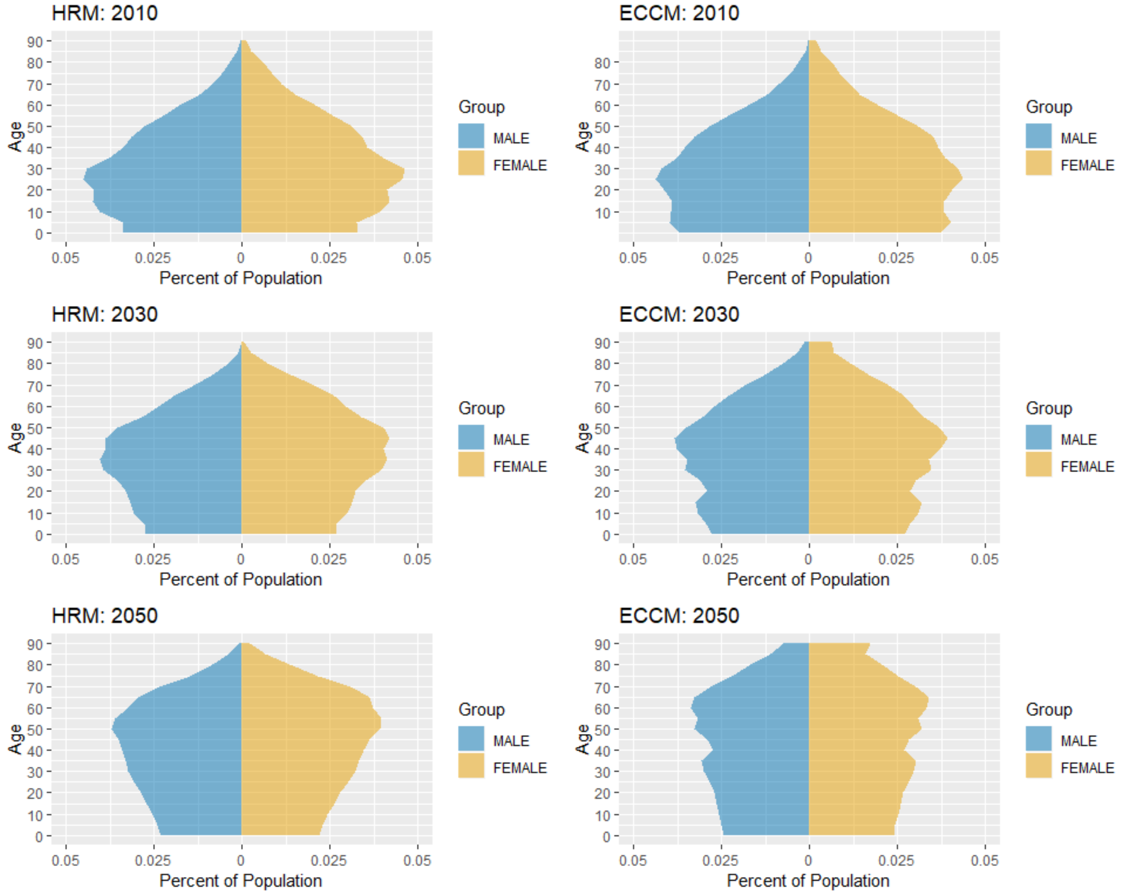
Source: Author's elaboration.

2.7 Comparing Population Projection and Household projection by number, size, and composition

Population Projection

It is through the population composition change that a demographic projection will determine the number, the size, and the composition of households. As seen before, the methods analyzed here have different approaches, while the classical approach uses the Cohort-Component Method, the second one uses the "Extended" Cohort-Component version of the same method. However, those methods did not present large differences. The graph below shows São Paulo's population pyramid projected from 2000 to 2010 and the projected population pyramid to 2050 for the constant rates scenario.

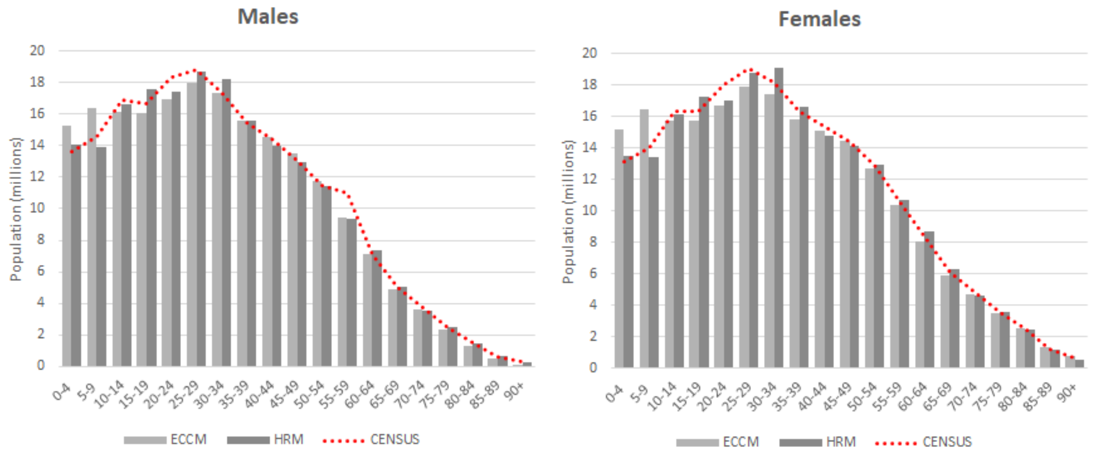
FIGURE 4 – São Paulo's population projection (2010-2050) by Headship Rate Method (HRM) and Extended Cohort-Component Method (ECCM): most likely scenario (2)



Source: Author's elaboration using softwares Popgroup and Profamy.

Over a 10-year horizon, comparing the 2000 to 2010 projections, both methods seem to be consistent with Census observations. The graph below shows the male and female age distribution. The respective relative errors compared to census observations stay within the reasonable margin of 10% of errors for subnational projections. The most concerning point that should be highlighted is the ECCM fertility estimation that provided two birth cohorts larger than Census observation, which can affect further household estimations.

FIGURE 5 – São Paulo's population projection (2000-2010) by Headship Rate Method (HRM) and Extended Cohort-Component Method (ECCM)

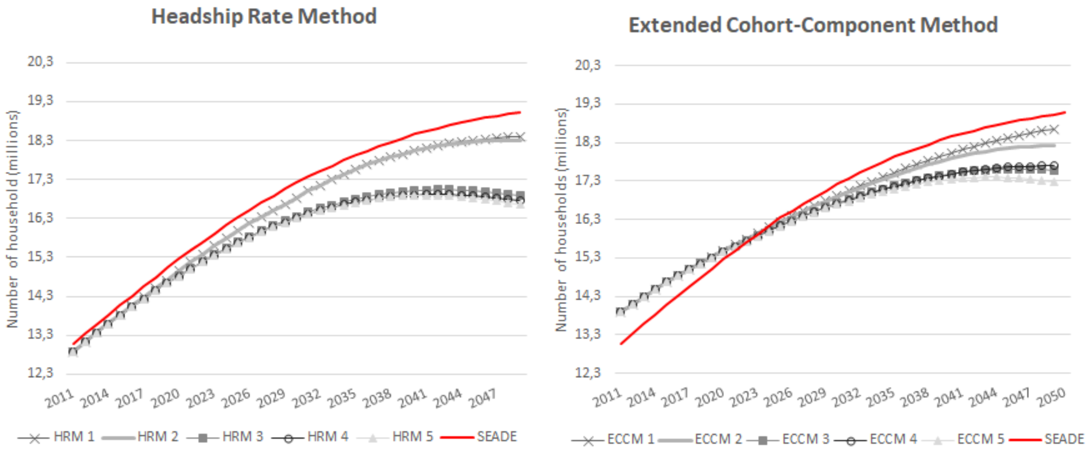


Source: Author's elaboration using softwares Popgroup and Profamy.

Total number of households

The first scenario which considers the life expectancy reaching its maximum combined with the TFR replacement level produces the higher number of households followed by the scenario which considers the maximum life expectancy but combined with low TFR. However, the total number of households of these scenarios is lower than projected by the SEADE Foundation during the entire projection period. The variation in life expectancy was more significant for the total number of households than the variation in TFR. The lowest household growth rate was obtained by scenario 4, which considers lower life expectancy combined with TFR replacement level. According to this scenario, there will be 29% more households in 2050 compared to 2010, while SEADE Foundation project 45% more households compared to 2010 (Table 18).

FIGURE 6 – Total number of households by Headship Rate Method (HRM) and Extended Cohort-Component Method (ECCM), São Paulo (2010-2050): different scenarios



Source: Author’s elaboration using software Popgroup and Profamy.

TABLE 18 – Household increase percentage from base population by Headship Rate Method (HRM) and Extended Cohort-Component Method (ECCM), São Paulo

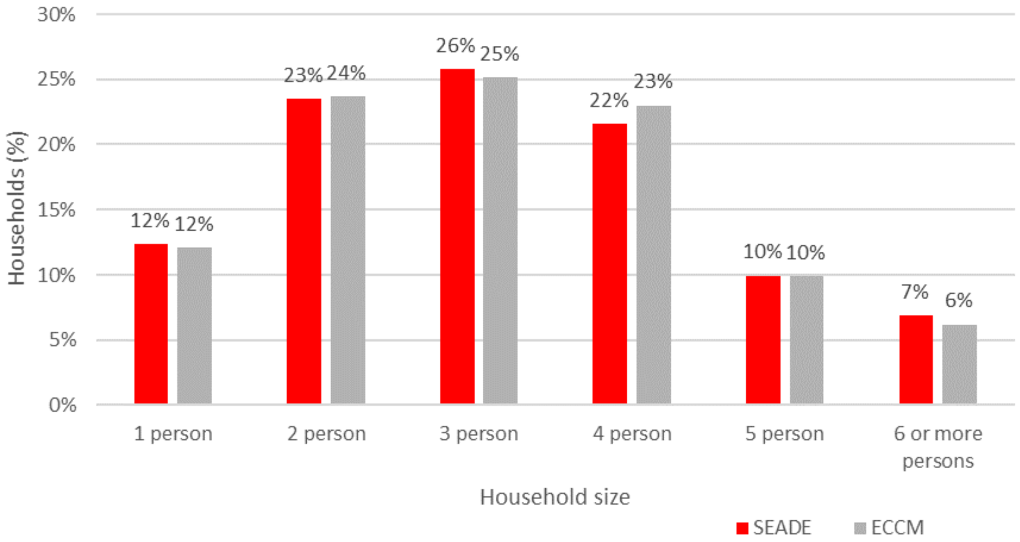
Household increase percentage	Headship Rate Method (HRM)	Extended Cohort-Component Method (ECCM)
1. Maximum e0 + Maximum TFR	43%	35%
2. Maximum e0 + Minimum TFR	41%	31%
3. 2010 Constant rates	31%	26%
4. Minimum e0 + Maximum TFR	30%	27%
5. Minimum e0 + Minimum TFR	29%	23%

Source: Self elaboration using software Popgroup and Profamy.

Household size

As pointed out before, one of the criticisms about the Headship Rate Method is that the usual approach of the method does not allow a household size projection. However, it is possible to overcome this problem by calculating the household size rates instead of headship rates. In this study, this approach will not be used. Meanwhile, the ECCM showed a consistent household size projection from 2000 to 2010 compared to the official SEADE Foundation estimation (Figure 7).

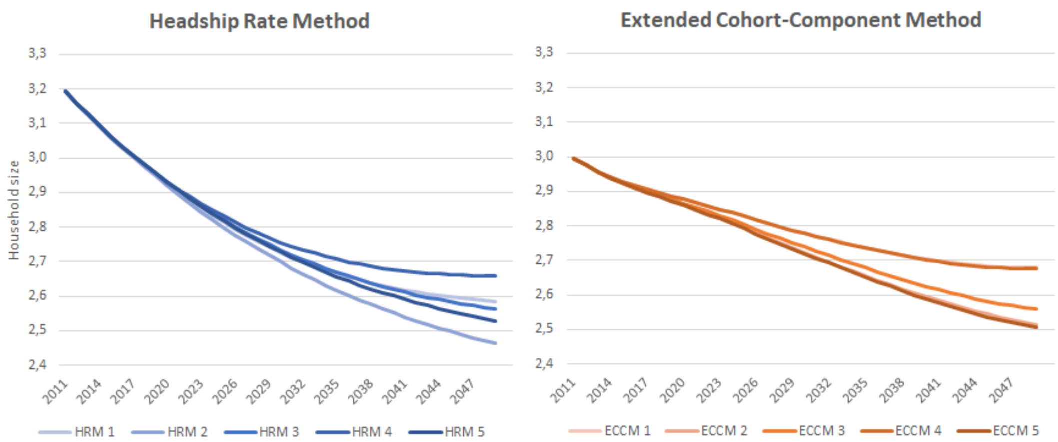
FIGURE 7 – São Paulo's household distribution by size in 2010, by Extended Cohort-Component Method (ECCM)



Source: Author's elaboration using SEADE Foundation (2018) and Profamy Software.

The average household size can be derived by division between the total projected population and the projected total number of households. The average household size was about 3 and 3.2 in early 2010 and it tended to decrease constantly in future decades. The different scenarios produced a range from 2.5 to 2.68 household average size in ECCM and from 2.46 to 2.6 average persons in HRM.

FIGURE 8 – São Paulo's household average size (2010-2050) by Headship Rate Method (HRM) and Extended Cohort-Component Method (ECCM): different scenarios



Source: Author's elaboration using software Popgroup and Profamy.

The Tables 19, 20 and 21 below summarize the differences among the two household projection methods in terms of percentage difference in population size, number of households and average household size.

TABLE 19 – Percentage Difference between the ProFamy (ECCM) and Headship Rate Method (HRM) projections (2000-2010) and the 2010 Census Observation for São Paulo State

São Paulo State (2010)	ProFamy (ECCM)	Headship Rates Method (HRM)
Population size	-0.35	-0.00
Number of households	0.96	-1.96
Average household size	0.00	-0.04
% of 1 person households	-1.78	-
% of 2-3 person households	-0.99	-
% of 4+ person households	-1.85	-

Source: Author's elaboration.
*Percentage difference = [(ProFamy projection - HRM projection) / HRM projection] × 100.

TABLE 20 – Percentage differences of household number and population size for São Paulo State between the Headship Rate Method projections and the ProFamy (ECCM) Projections*

São Paulo State	2020	2030	2040	2050
Number of Households				
Scenario 1	3.6	0.6	-0.1	1.4
Scenario 2	3.6	0.4	-0.8	-0.4
Scenario 3	4.2	2.3	2.8	4.0
Scenario 4	4.2	2.5	3.4	5.6
Scenario 5	4.2	2.3	2.7	3.7
Population Size				
Scenario 1	2.0	2.4	3.0	5.1
Scenario 2	1.4	1.2	0.8	1.7
Scenario 3	2.0	2.5	3.0	4.0
Scenario 4	2.4	3.3	4.4	6.5
Scenario 5	1.7	2.0	2.2	2.9

Source: Author's elaboration.
*Percentage difference = [(ProFamy projection - HRM projection) / HRM projection] × 100.

TABLE 21 – Population size, numbers of households and household size in São Paulo State considering the most likely scenario by projection method, 2010 to 2050

Year	2010	2020	2030	2040	2050	Percentage Change (2050 vs 2010)
Population size (Millions)						
ProFamy	41,26	44,29	45,98	46,39	45,82	+11.05%
HRM	41,26	43,66	45,44	46,02	45,04	+ 9.15%
Number of Households						
ProFamy	12,61	15,48	16,89	17,88	18,21	+44.43%
HRM	12,61	14,94	16,82	18,02	18,28	+44.96%
Household average size						
ProFamy	3.27	2.86	2.72	2.59	2.51	-29.15%
HRM	3.27	2.92	2.70	2.55	2.46	-24.70%

Source: Author's elaboration.
* Percentage change = [(Number in 2050-Number in 2010) / Number in 2010] × 100.

Household type

When comparing results by household composition between ECCM and HRM, one important methodological issue came to light. As explained before, HRM household type classification (living alone, couple with children, couple without children, mono-parental, extended household) is based on reference person question, present in Census and PNAD questionnaires. This is also how Demography literature have been dealing in cross-sectional context countries. However, instead of using Census reference person, the ECCM uses its own approach to follow an "ego" to compose the households, which is the female adult always when it is possible, as Zeng et al. (2014b) argue:

> We follow Brass' marker approach to identify households based on individuals' characteristics. Brass (1983) calls the reference person a household "marker". In Brass's original work, the nuclear family-status life table models developed by Bongaarts (1987), and the general family-status life table model including nuclear and three-generation families of Zeng (1986, 1988, 1991a), only female adults are chosen as markers, which implies a female-dominant one-sex model. In the ProFamy model developed in this book, both sexes are included; a female adult, or a male adult when a female adult is not available, is identified as the reference person (or "marker") of the household (ZENG et al., 2014b, p. 22).

The household distribution based on this ego marker is called by Zeng et al. (2014b) as "Model-count", while the household distribution based on Census reference person is called "Direct-count". As the reference person and Brass' marker has different definition, and by consequence produce different households, the comparison between "Model-count" and "Direct-count" is only possible for the total number of households, household size and living alone household type, when the "ego" and the refence person are the same. These are the measures, that can be compared to the Census data and are used by Zeng et al. (2014b) to validate the model.

In conclusion, it is possible to derive all results by household type from ECCM “Model-count” using Brazilian data, however they have the disadvantage of not being comparable with cross-sectional surveys and Census household type based on reference person. No additional information is available to reproduce ECCM ego’s using Census data and this can represent an important barrier to official statistics offices. Further studies should be done comparing specific ECCM household composition types with Census data. Meanwhile, this work will focus on how we can improve the classic Headship Rate household projection approach.

CHAPTER 3 – IMPROVING THE CLASSIC HEADSHIP RATE HOUSEHOLD PROJECTION APPROACH: THE FUNCTIONAL PRINCIPAL COMPONENTS (FPC) MODEL

The usage of private software for population and household projection has many disadvantages. The user cannot control the entire projection process, such as manipulate the dataset or plot their own results. Furthermore, the user cannot choose the mortality, fertility or migration forecast methods, even though there is a more suitable model for their data. The previous ECCM and HRM household projections uses relational parameterized models. For mortality it uses “model life Tables” (e.g., COALE; DEMENY; VAUGHAN, 1983; UNITED NATIONS, 1982) and “Brass logit relational life table model” (e.g., MURRAY et al., 2003), while for fertility they are based on “Brass relational Gompertz fertility model” (BRASS, 1974), and other parameterized models (e.g., COALE; TRUSSELL, 1974; ROGERS, 1986) in population projections and estimations.

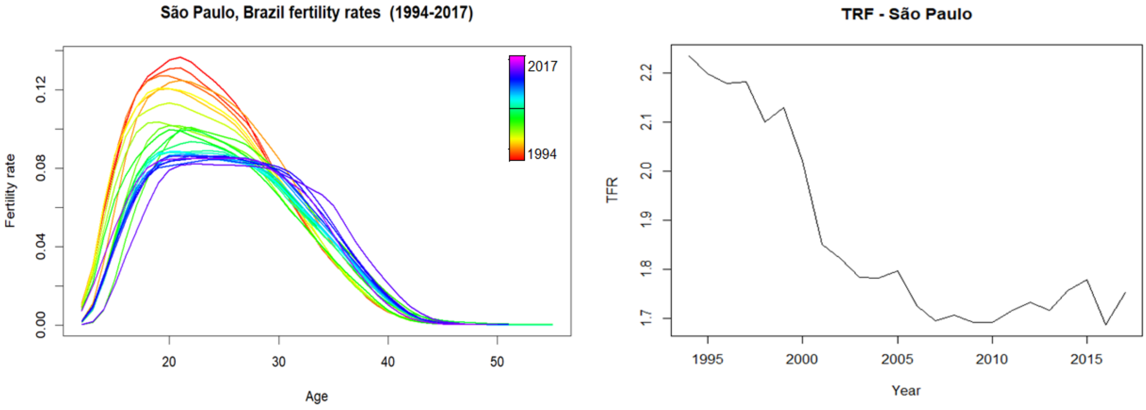
Here another Cohort-Component projection will be used to support the São Paulo household projection. The main idea is to reproduce a household projection in an open-source software and choose more recent fertility and mortality forecast models. This new approach will use stochastic forecasting of mortality and fertility. Instead of the scenario-based approach, which is considered flawed by the recent literature, the confidence intervals of the stochastic forecasting were considered to build these scenarios.

Moreover, a new alternative approach will be proposed for headship rates forecasting. As most of the classical Headship Rate projections consider the headship rates constant over time or creates scenarios without mathematical model basis, here it will be constructed a headship rate forecast by sex and household type based on the Lee-Carter method (LEE; CARTER, 1992).

3.1 Functional Principal Components (FPC) fertility forecasting

The fertility data source is the Brazilian Birth Information System (SINASC) which can provide the information of births by single year of age and single year since 1994. The raw data was smoothed with Weighted Penalized Regression Splines function (BOOTH; HYNDMAN; TICKLE, 2014). While the population data source is from IMP/SEADE Foundation which provides an estimation of the population size by 5 age-group, sex, and single year. The interpolation Sprague-Karup method (DE KERF, 1975) to disaggregate the 5 age-group was applied in order to get the single age population counts. The smoothed age-sex specific fertility rates are shown in the graph below.

FIGURE 9 – Age-Specific Fertility Rates (ASFR) and Total Fertility Rate (TFR), São Paulo, Brazil (1994-2017)



Source: Source: Author's elaboration using SINASC data.

The age-specific fertility rates have been changing since 1994 both in terms of level and shifting. In 1994 the modal age of fertility was located in the early twenties while in 2017 the modal age of fertility is located closer to 30 years. That change leads to a drop in Total Fertility Rate from 2.2 in 1995 to around 1.7 in 2017 (Figure 9).

The functional data analysis technique based on Ramsay and Silverman (2005) and Hyndman and Ullah (2007), is a nonparametric method for modeling and forecasting fertility rates and log mortality rates. According to Booth; Hyndman and Tickle (2014), this approach extends the Lee-Carter method and can be described in three ways: 1. the log mortality rates are smoothed prior to modeling; 2. functional principal components analysis is used; 3. more than one principal component is used in forecasting;

The functional principal components analysis, also known as Hyndman-Ullah (HU) method can be considered a successor to the Lee-Carter method and can provide robust forecast age-specific fertility rates. Given $z_t(x)$, the log of the observed fertility rate for age x in year t, we assume there is an underlying smooth fertility function $f_t(x)$ that we are observing with some error:

$$z_t(x_i) = f_t(x_i) + \sigma_t(x_i)\varepsilon_{t,i} \quad (6)$$

Where $\sigma_t(x_i)$ allows the amount of noise to vary with x_i in year t, thus rectifying the assumption of homoscedastic error in the Lee Carter model. The age-specific fertility curves are decomposed into orthogonal functional principal components and their uncorrelated principal component scores:

$$f_t(x_i) = a(x) + \sum_{j=1}^{J} b_j(x)k_{t,j} + e_t(x), \quad i = 1,\ldots,p, t = 1,\ldots,n \quad (7)$$

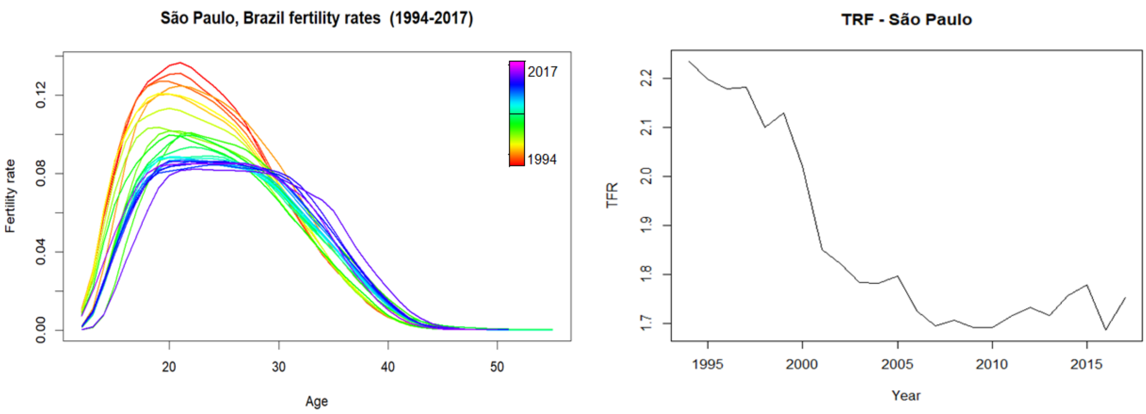
Where a(x) is the mean function estimated by $\hat{a}(x) = \frac{1}{n}\sum_{t=1}^{n} f_t(x_i)$; $\{b_1(x), \dots, b_j(x)\}$ is a set of the first J functional principal components; $\{k_{t,1}, \dots, k_{t,j}\}$ is a set of uncorrelated principal component scores; $e_t(x)$ is the residual function with mean zero; and J < n is the number principal components used. The h-step-ahead forecast of z_{n+h}(x) can be obtained by:

$$\hat{z}_{n+h|n}(x) = E\left[z_{n+h|n}(x)|\boldsymbol{L}, \boldsymbol{B}\right] = \hat{a}(x) + \sum_{j-1}^{j} b_j(x)\hat{k}_{n+h|n,j} \quad (8)$$

Where **L** $= \{z_1(x), \dots, z_n(x)\}$ is the observed fertility data, $\boldsymbol{B} = \{b_1(x), \dots, b_j(x)\}$ is the set of functional principal components, $\hat{k}_{n+h|n,j}$ denotes the h-step-ahead forecast of $k_{t,j}$ using a univariate time series model, such as the optimal ARIMA model chosen by 'forecast' R package. More details can be found in Hyndman and Ullah (2007) or Booth; Hyndman and Tickle (2014). The model can be found implemented in the R package called "Demography".

The observed fertility data (grey color) for São Paulo state and the forecast fertility data (rainbow colored) are shown in the Figure 10. The model captured the decreasing fertility trend for younger ages (from 25 to 30 years old), while for older ages there is a rapid increase showing fertility aging and postponement. This change in the fertility curve will represent a change in TFR show in the graph below. According to the functional principal components model, the TFR which was 1.75 in 2017 for São Paulo, will show a small recovering reaching 1.88 in 2030, considering a 95% confidence interval from 1.21 to 2.86.

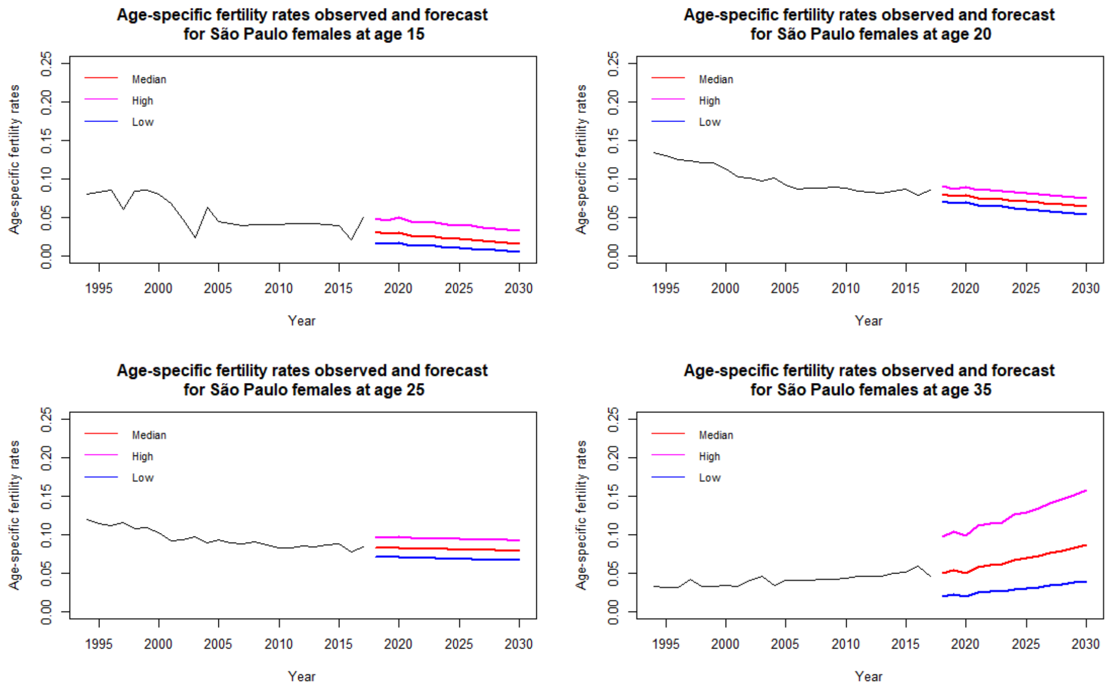
FIGURE 10 – Age-Specific Fertility (ASFR) rates and Total Fertility Rate (TFR) forecast, São Paulo, Brazil (2017-2030)



Source: Author's elaboration using SINASC data.

It also possible to analyze the age-specific fertility rate for some selected ages (15, 20, 25 and 35) and see how the ASFR behaves in the period. For almost all ages the trend is monotonically decreasing, and the functional principal components model produces a reasonable expected future trend with tight 95% confidence intervals. While for older ages, the trend is increasing over time and the functional principal components model produces larger confidence intervals (Figure 11).

FIGURE 11 – Age-Specific Fertility Rates forecast by age 15, 20, 25 and 35. São Paulo, Brazil (2017-2030)

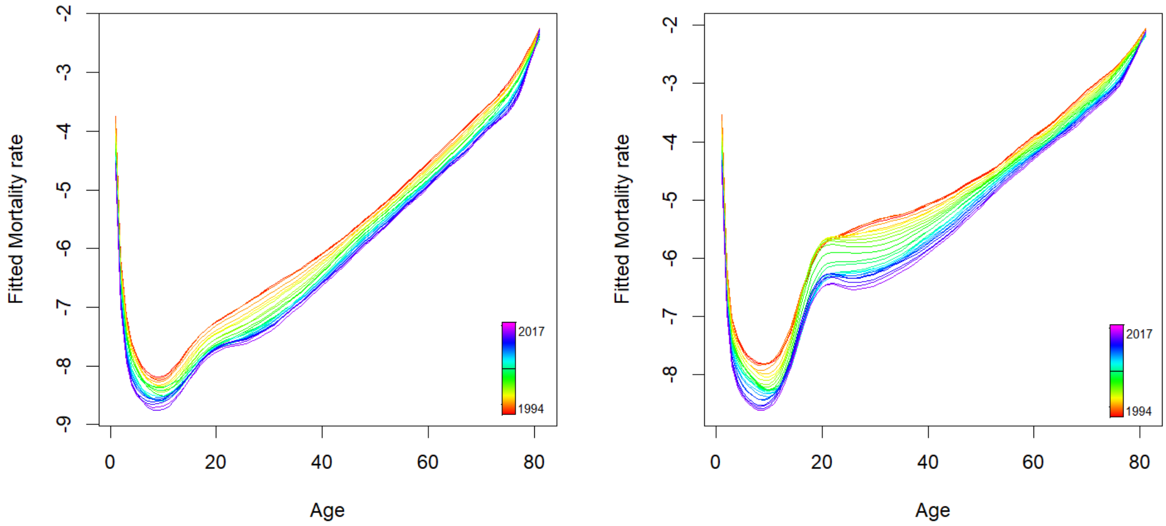


Source: Author’s elaboration using SINASC data.

3.2 Functional Principal Components (FPC) mortality forecasting

The data source for mortality counts was obtained from the Brazilian Mortality Information System (SIM) which provides the number of deaths for single years of age and single years for São Paulo state. As fertility data, the raw data was smoothed with a Weighted Penalized Regression Splines function (BOOTH; HYNDMAN; TICKLE, 2014), where the observational error has higher variance at older ages (when the populations are small) and at young ages (when the mortality rates are small). While the population information was obtained by the IMP/SEADE Foundation that provides the counts of population by single year of time and age-groups. The Sprague-Karup interpolation method (DE KERF, 1975) was applied to disaggregate the 5 age-group in order to get the single age population counts. The smoothed age-sex-specific mortality rates are shown on a log scale in the Figure 12.

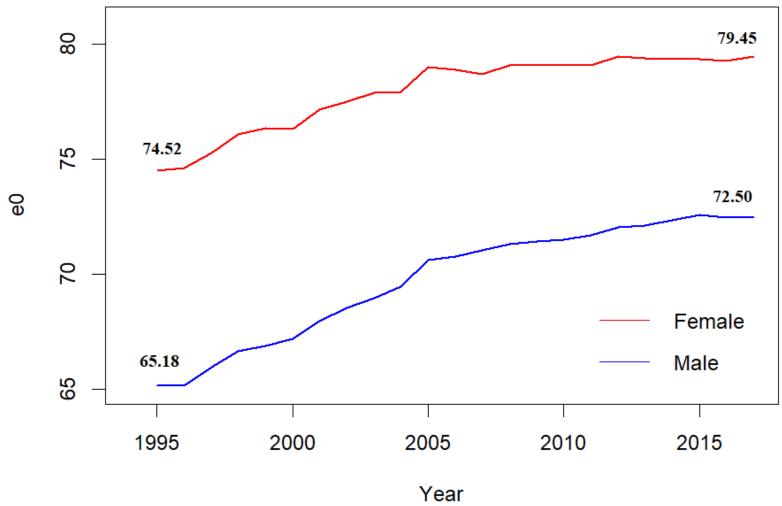
FIGURE 12 – Age-Specific Mortality Rates (ASMR) by age and sex (women – left, men – right). São Paulo, Brazil (1994-2017)



Source: Author's elaboration using SIM data.

As a consequence of the drop of mortality rates since 1994 for all ages in the São Paulo population, the life expectancy at birth has been increasing. The gain for females (from 74.52 in 1995 to 79.45 in 2017) was higher compared to males (from 65.18 in 2015 to 72.50 in 2017), and the sex gap in life expectancy showed a small decrease (from 9.34 in 1995 to 6.95 in 2017).

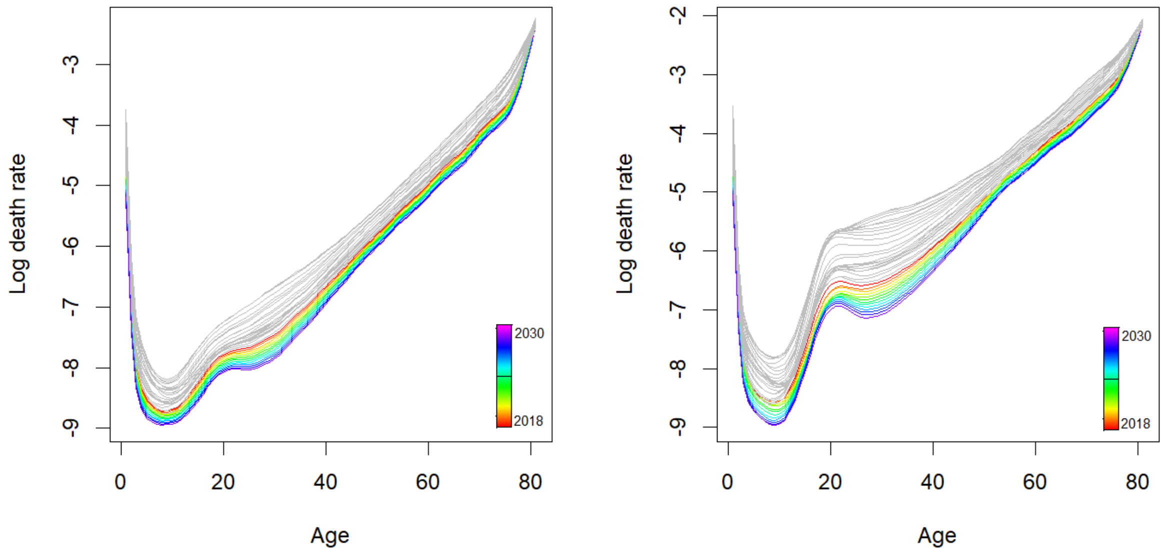
FIGURE 13 – Observed life expectancy at birth by sex. São Paulo, Brazil (1994-2017)



Source: Author's elaboration using SIM data.

The functional principal components analysis can also provide us with a robust forecast age-specific mortality rates in the same way as fertility model described before. The observed mortality data (grey color) for São Paulo state and the forecast mortality data (rainbow colored) are shown in the Figure 14.

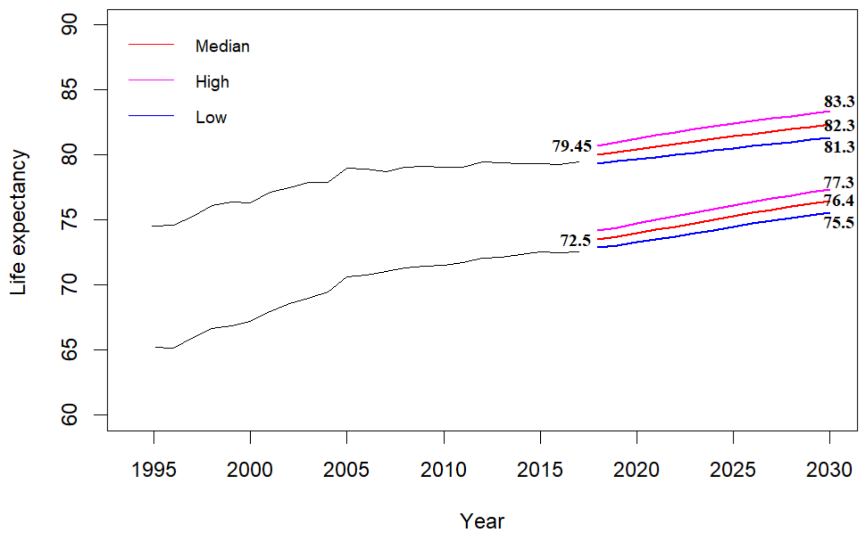
FIGURE 14 – Age-Specific Mortality Rates (ASMR) forecast by sex, São Paulo, Brazil (2018-2030)



Source: Author's elaboration using SIM data.

According to the functional principal components model, the age-specific mortality rates will result in increasing life expectancy at birth for both sexes. The graph below shows the 95% confidence interval for life expectancy at birth. For females, a life expectancy of 82.3 years towards 2030 is expected (in a 95% confidence interval of 81.2 to 83.3 years) while for males a life expectancy of 76.4 years towards 2030 is expected (in a 95% confidence interval of 75.5 to 77.3 years.

FIGURE 15 – Observed and forecast life expectancy at birth by sex, São Paulo, Brazil (2018-2030)



Source: Author's elaboration using SIM data.

Some selected age-specific mortality rates were chosen in order to see how the ASMR behaves in the period. For almost all ages, the male mortality trend is monotonically decreasing, and the functional principal components model provides a reasonable expected future trend with tight 95% confidence intervals. The ASMR for ages 0, 25, 60 and 70 are shown in the graph below.

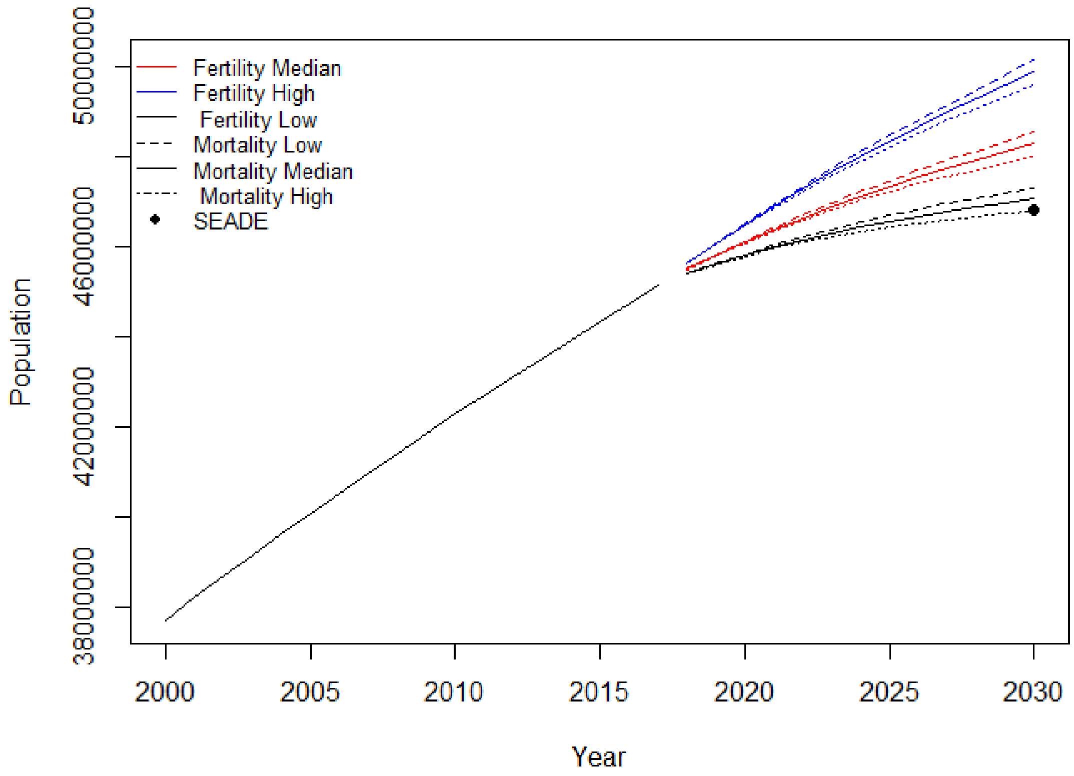
FIGURE 16 – Age-Specific Mortality Rates forecast by age 0, 25, 60 and 70. São Paulo, Brazil (2018-2030)

Source: Author’s elaboration using SIM data.

3.3 Cohort-component population projection

Once the age-specific fertility rates and the age-specific mortality rates are forecast, the Leslie matrix for Cohort-component projection can be calculated for each forecast year for the São Paulo population. Combining the low and high 95% confidence interval boundaries in the mortality and fertility forecasts, it produced 9 scenarios for São Paulo’s population growth. Comparing with SEADE Foundation projection which estimate a total population size of 46,825,450 for São Paulo state in 2030, the MedMin scenario of FPC method estimated a total population of 48,245,739 resulting a percentual difference of 3,03%, while the scenario MaxMin estimated 46,790,138 resulting a percentual difference of 0,08% from SEADE Foundation projection (Figure 17). Figure 18 shows the MedMin scenario age distribution which was withing the 10% from SEADE Foundation age distribution, differing only in open age group 80+. The differences among total population projection by different methods are presented in Table 22.

FIGURE 17 – São Paulo Total Population forecast by scenarios (2018-2030)



Source: Author’s elaboration.

FIGURE 18 – Comparison between FPC MedMed scenario age distribution and SEADE Foundation age distribution (2030)

Source: Author's elaboration.

TABLE 22 – Differences among total population projection by different methods, São Paulo state, Brazil, 2030

Number of households by type	Total pop. 2030	Diff (#)	Diff (%) from SEADE Foundation
FPC	48,245,739	-1,420,289	3,03%
ECCM (PROFAMY)	45,482,993	1,342,457	2,87%
HRM (POPGROUP)	43,882,462	2,942,988	6,29%
SEADE Foundation	46,825,450	-	-

Source: Author's elaboration. SEADE Foundation (2017).
*Diff = [(Projection – 2010 Census) / 2010 Census] × 100.

3.4 A new approach for Headship Rate forecasting

Most of the Brazilian experience using household projection via the Headship Rate Method assumes that the headship rates will be constant over time, which is not true. In fact, when considering the total population headship rates, it can have small changes over time. However, when we look at headship rates by sex and/or household type they have some significant changes. In Brazil, males experienced a great headship rate decrease in the period 2000-2010, while for females there was a marked increase (Figure 46, Appendix). However, as discussed before, we should be aware that the question about who the reference person inside the household in Brazilian Census is has changed from 2000 to 2010. Consequently, there is an uncertainty about why the female's headship rates had such a rapid increased compared to male rates:

1. It reflects a cultural change linked to gender relations, which changes the understanding of the question. In this case, there is no real increase in female-headed households, just the same household that self-declares differently from the previous census.

2. There is indeed a growth of female headed households. In this case, a new household is created by separation, leaving the parents' house, widowhood, or any other reason, which is headed by a woman.

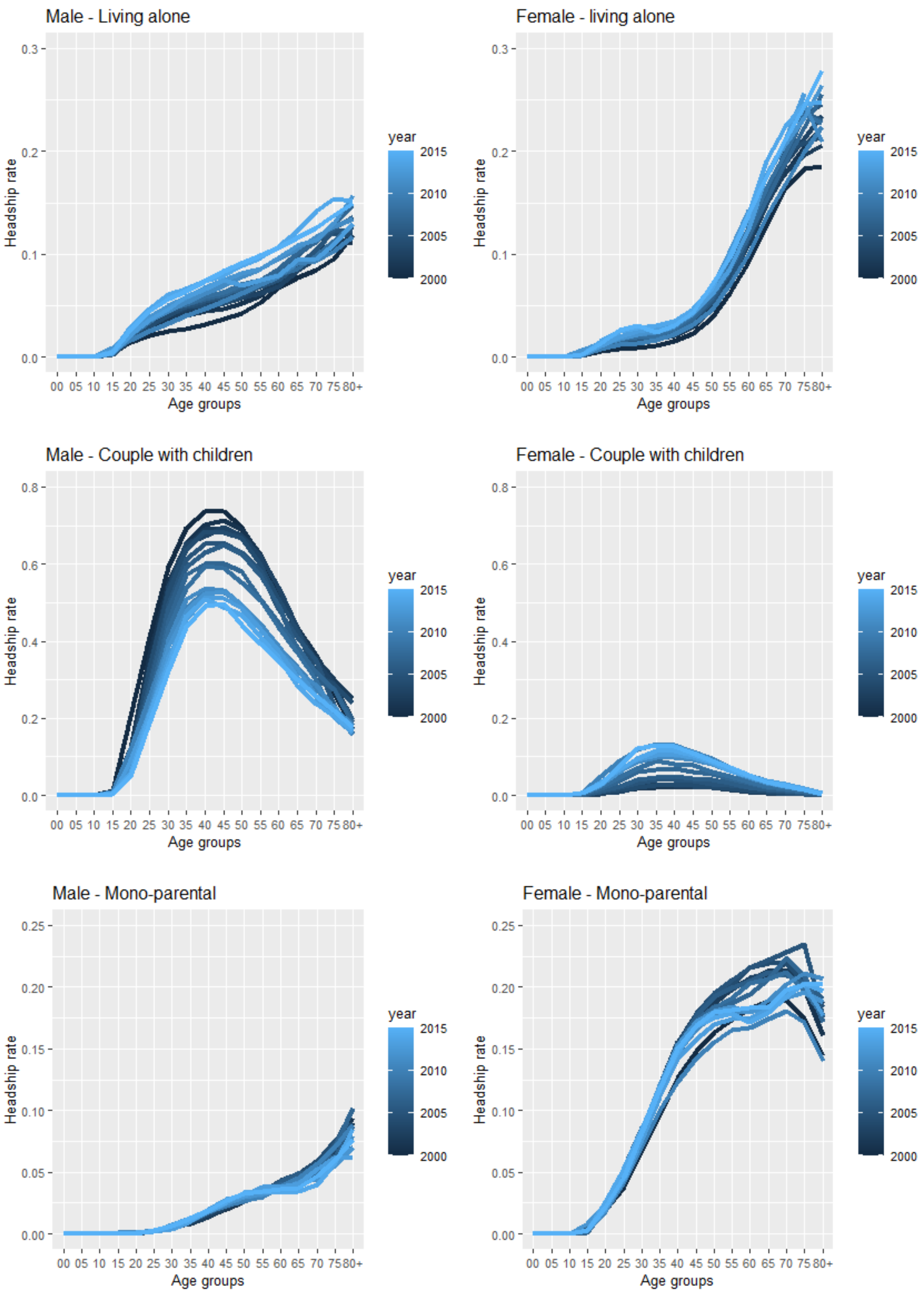
3. A mix of both effects.

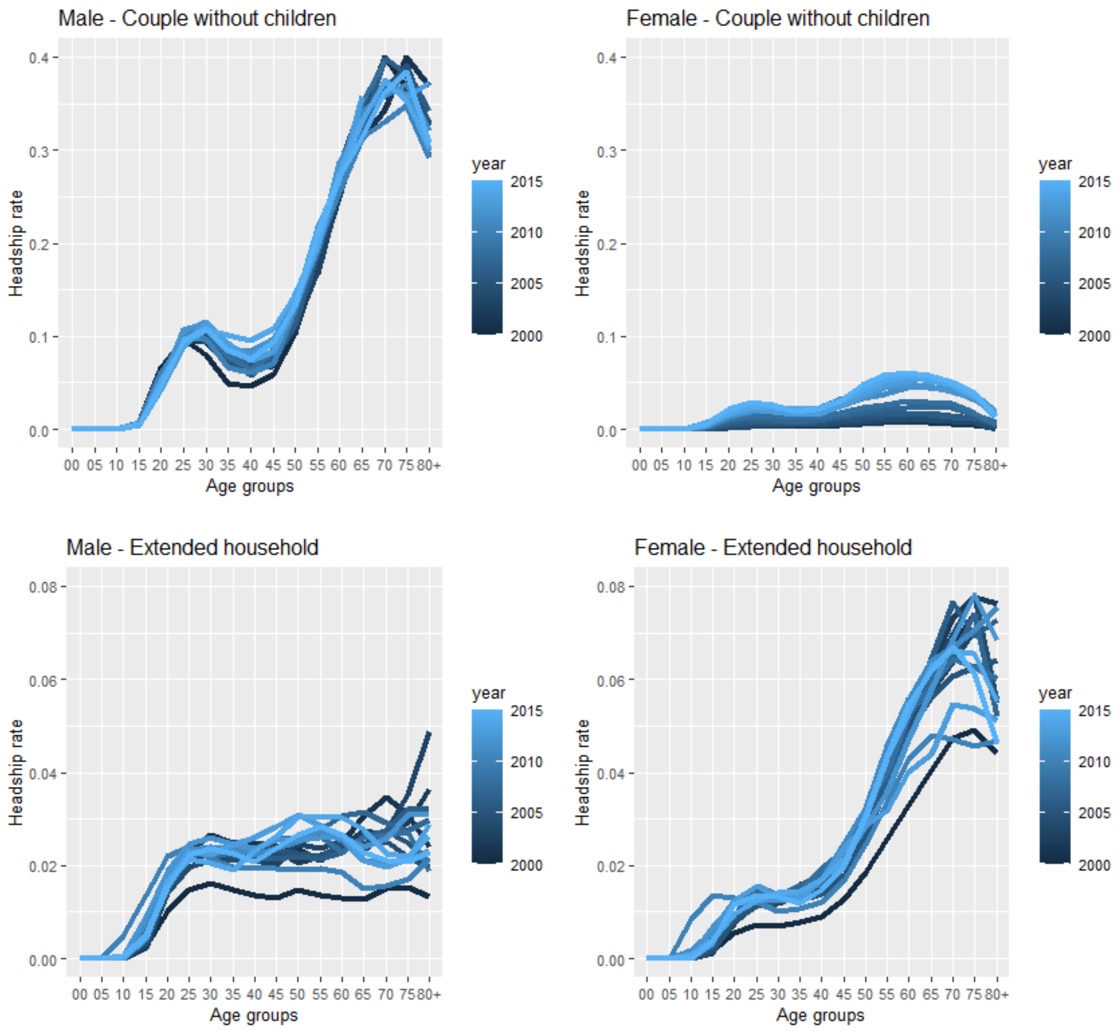
To solve this problem, the headship rates calculated here were taken from the Brazilian Household National Sample Survey (PNAD). Unlike the Demographic Census, the PNADs have maintained the same head identification since 1992 until 2015, which avoids the interpretation change effect.

> […] For PNAD, which began in 1967 and had annual uninterrupted collection from 1976 (except in the census years), the term "head" was used until the 1990 survey. In 1992, the first survey after census of the decade, unlike the one that adopted the term "responsible person", PNAD adopted the term "reference person". This was used until the last survey of the series without any changes in 2015 (CAVENAGHI; ALVES, 2018, p. 49).

The headship rates were calculated for PNADs from 2001 to 2015 and 2000 and 2010 Census. However, as PNAD is a sample survey and has a much smaller population than the Census, when the headship rate is disaggregated by single age, sex and household type there is a great variation and some gaps in the data, even for a big Brazilian State like São Paulo. Because of that, it was necessary to gather more data and the headship rates were calculated for the South East Region as a whole, including 3 more states (Rio de Janeiro, Minas Gerais, and Espírito Santo). The graph 17 show the headship rates for each sex and household type from 2000 to 2015 in the Brazilian South East Region.

FIGURE 19 – Headship rates by household type, Brazilian South East (2000-2015)





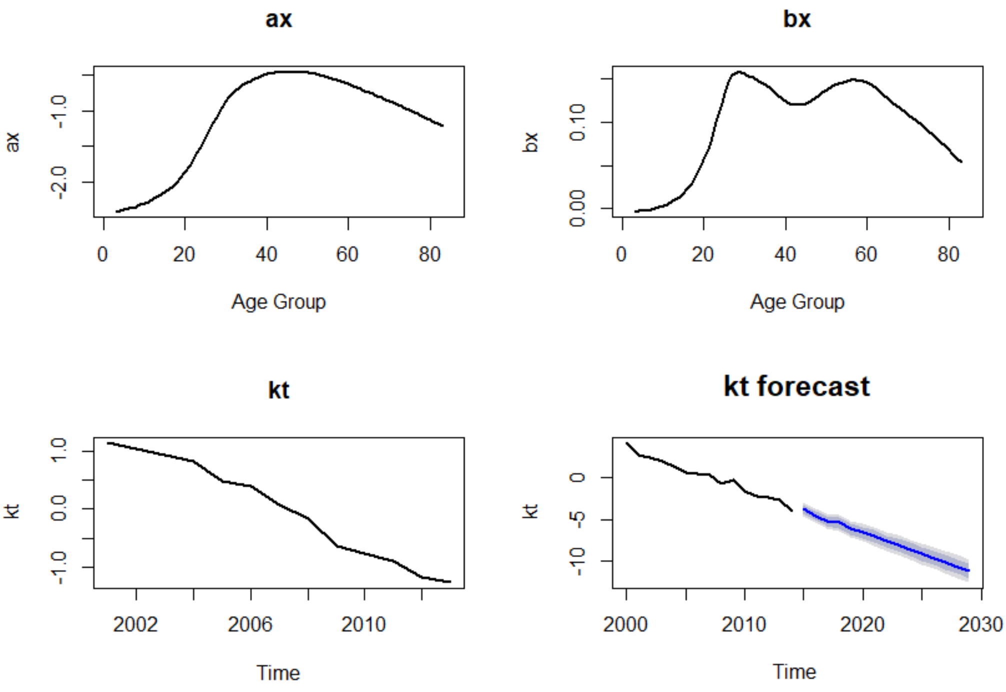
Source: Author's elaboration.

The model for projecting headship rates proposed here follows a Lee-Carter approach, which uses the SVD decomposition to separate the log of a defined rate into three effects: a general mean pattern, an effect of change at each age and other effect of change over time. The rate here, instead of the mortality rate, are the headship rates calculated above. Then, the model can be written as:

$$\ln[h(x,t)] = a(x) + b(x)k(x) + e(x,t) \qquad (9)$$

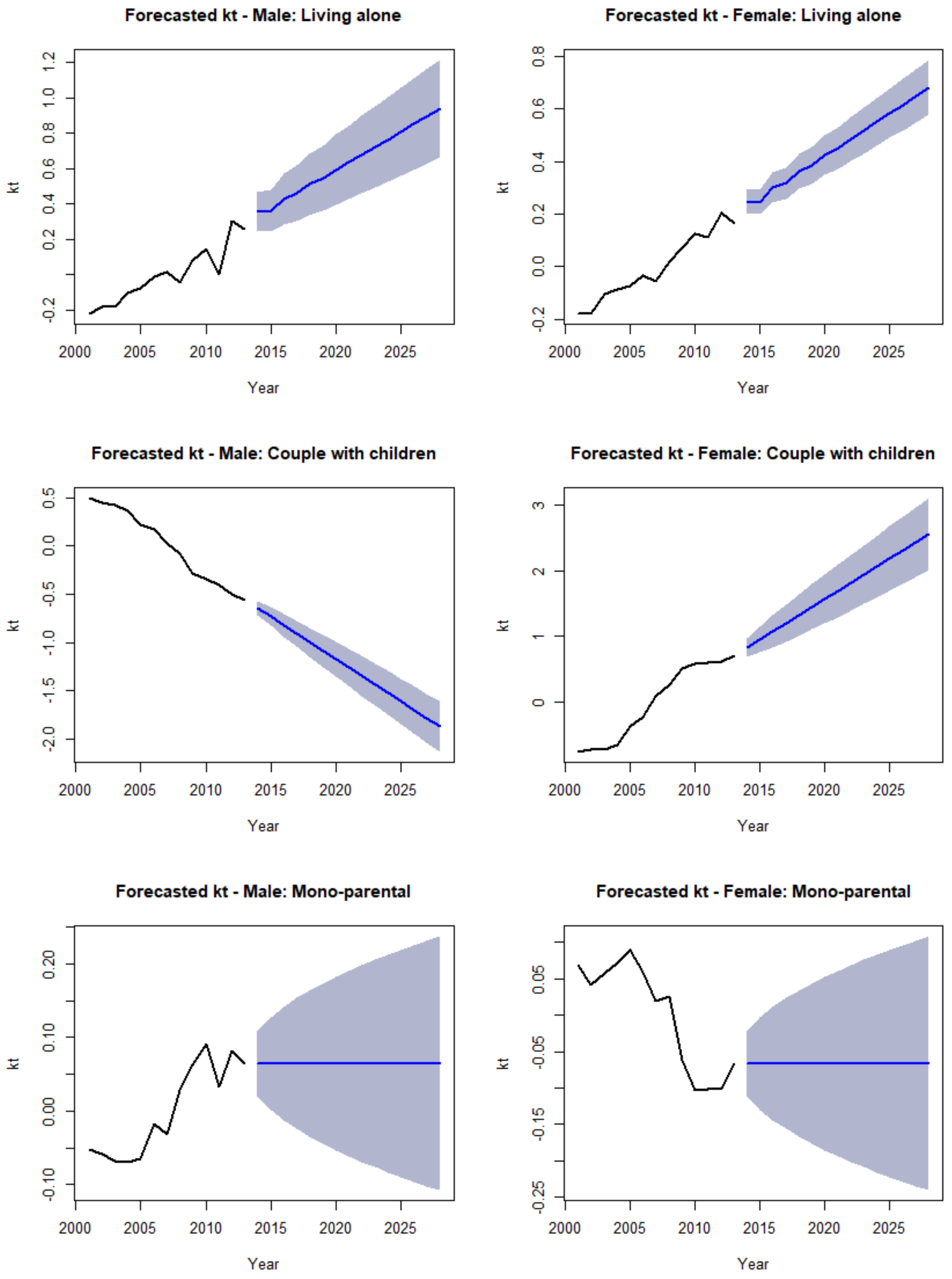
Where $h(x,t)$ is the headship rate at age x in year t; $k(t)$ is the level of headship rate over time; $a(x)$ is the general pattern of headship rate by age, $b(x)$ is the relative speed of headship rate change at each age and $e(x,t)$ the residual error at age x and time t, with Normal distribution $(0,\sigma^2)$. To illustrate the procedure, the Figure 20 shows the decomposition effects ($a(x)$, $b(x)$ and $k(t)$) for Brazilian South East males living in household of a couple with children. Using ARIMA models chosen from 'forecast' R package, it possible to forecast the trend of the time effect (t) for h = 13 years towards 2030. In this case, the trend is clearly linear decreasing.

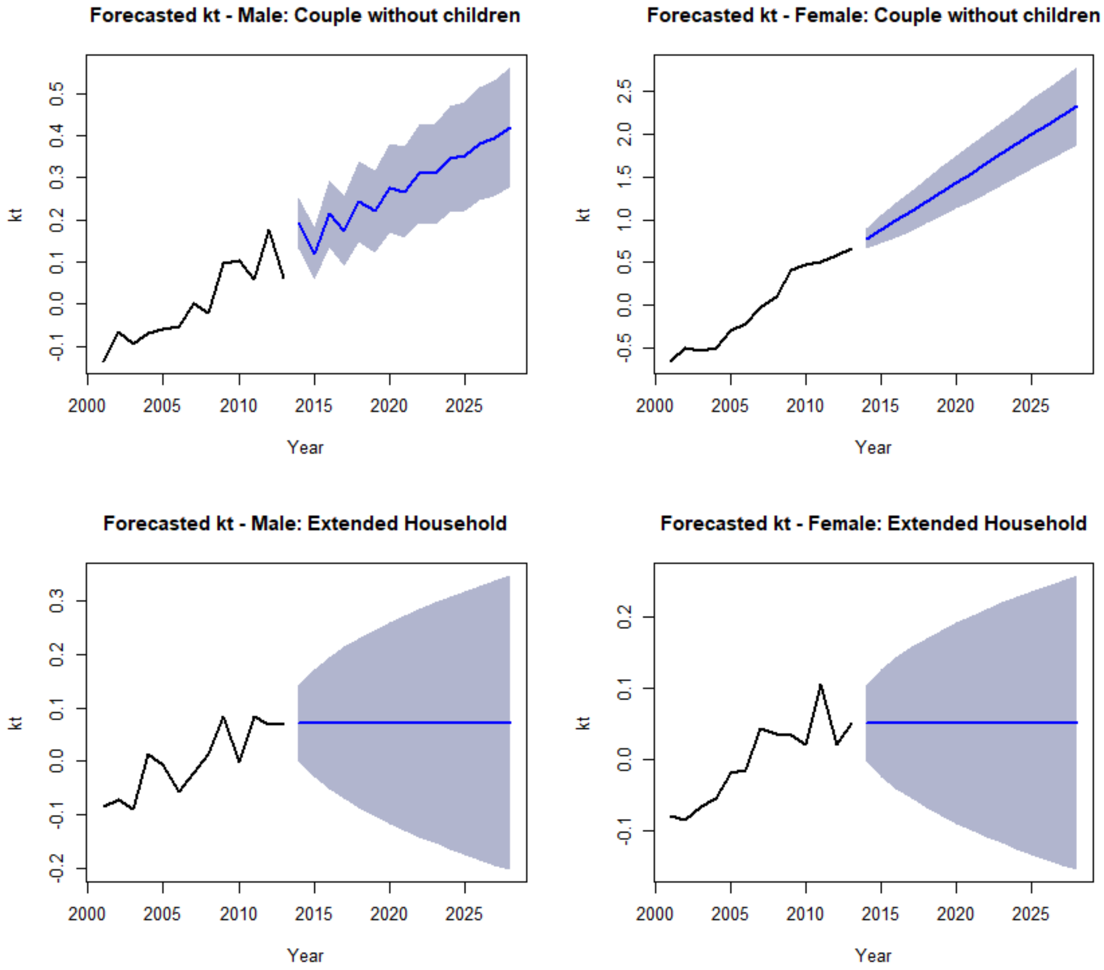
FIGURE 20 – a(x), b(x), k(t) effects and k(t) forecast for Brazilian South East Region males living in household of a couple with children



Source: Author’s elaboration.

FIGURE 21 – Time effect ($\boldsymbol{k_t}$) forecast by sex and household type, Brazilian South East (2015-2030)



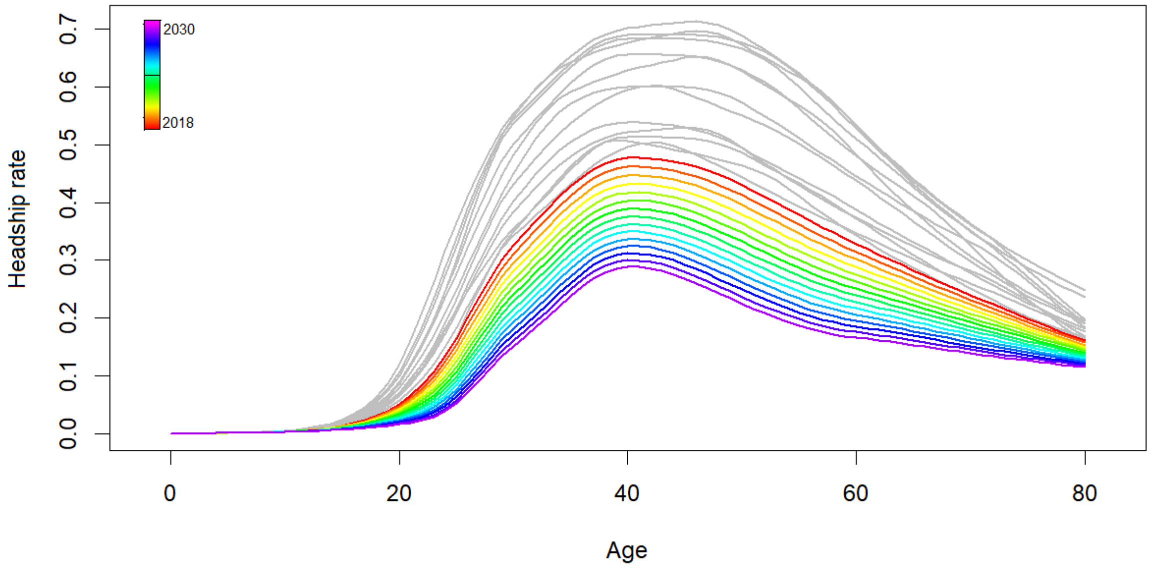


Source: Author's elaboration.

After obtained $a(x)$, $b(x)$ and forecast $k(t)$, we can estimate future headship rates ($h(x,t+h)$) as showing in the Figure 21. Because of the linear decreasing trend of $k(t)$, the future headship rates for males living in the household "couple with children" will have the same shape ($b(x)$ is constant in time) and each year will be smaller. It's important to note that, if (t) keep this linear trend constant over time, the headship rate will tend to zero. In a more distant forecast horizon, an asymptote (and consequently, new assumptions about this asymptote) would be necessary to restrict such trend and avoid loss of consistency. However, as the forecast horizon is reasonable short (13 years), it will be assumed that the trend observed during 2000-2015 will continue without any restriction.

Another strong assumption that should be discussed is related to sex and household type consistency. Considering one Lee-Carter model for each household type and sex yields independent results for each category. For example, a decreasing headship rate forecast for male living in "couple with children" household type will necessarily increase others male's household types. Also, it will directly increase female's headship rate forecast of the same household type. Because of that, the household number distribution resulting from this approach should be interpreted independently by sex and household type.

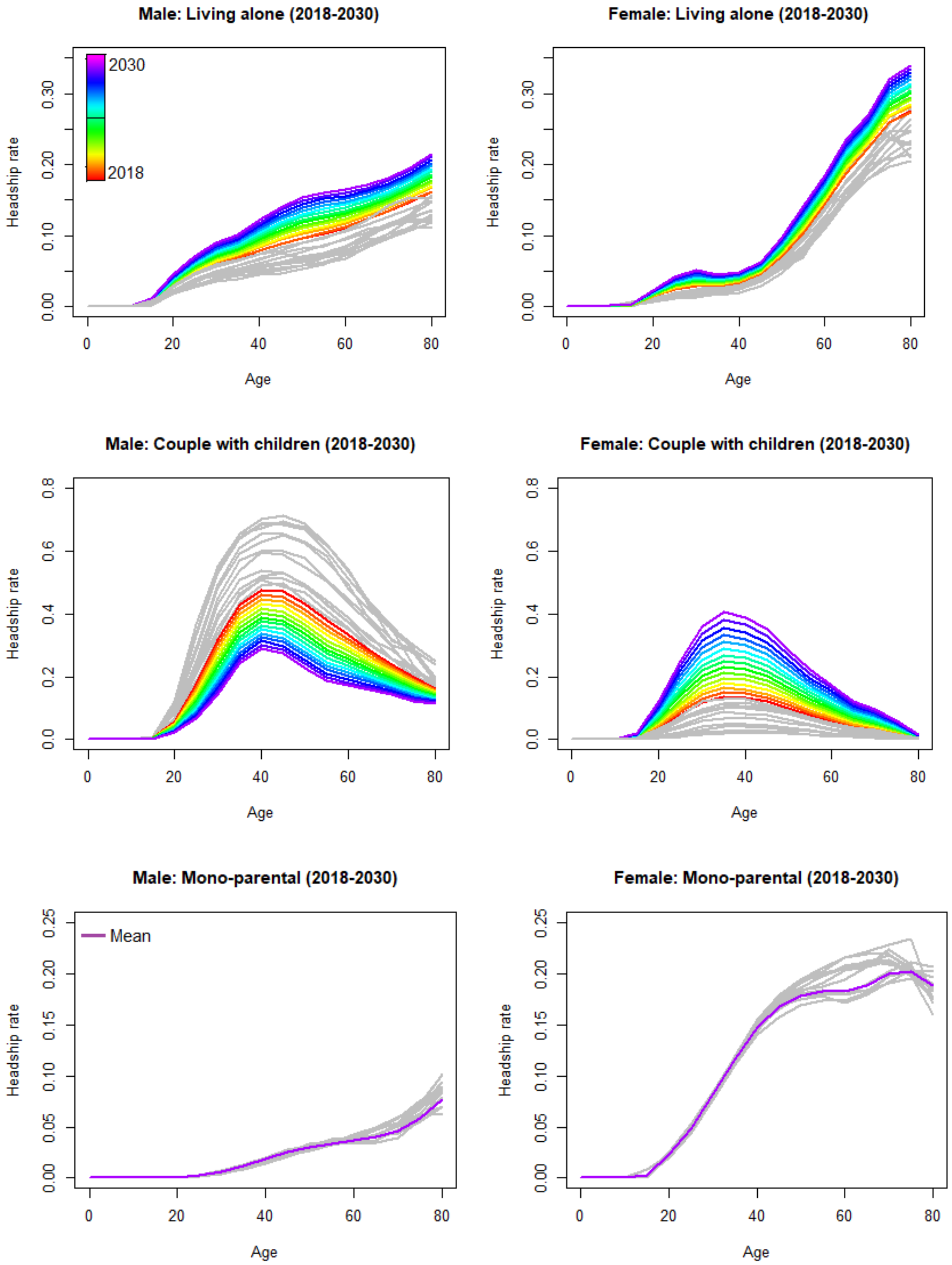
FIGURE 22 – Headship rate forecast, male: couple with children (2018-2030)

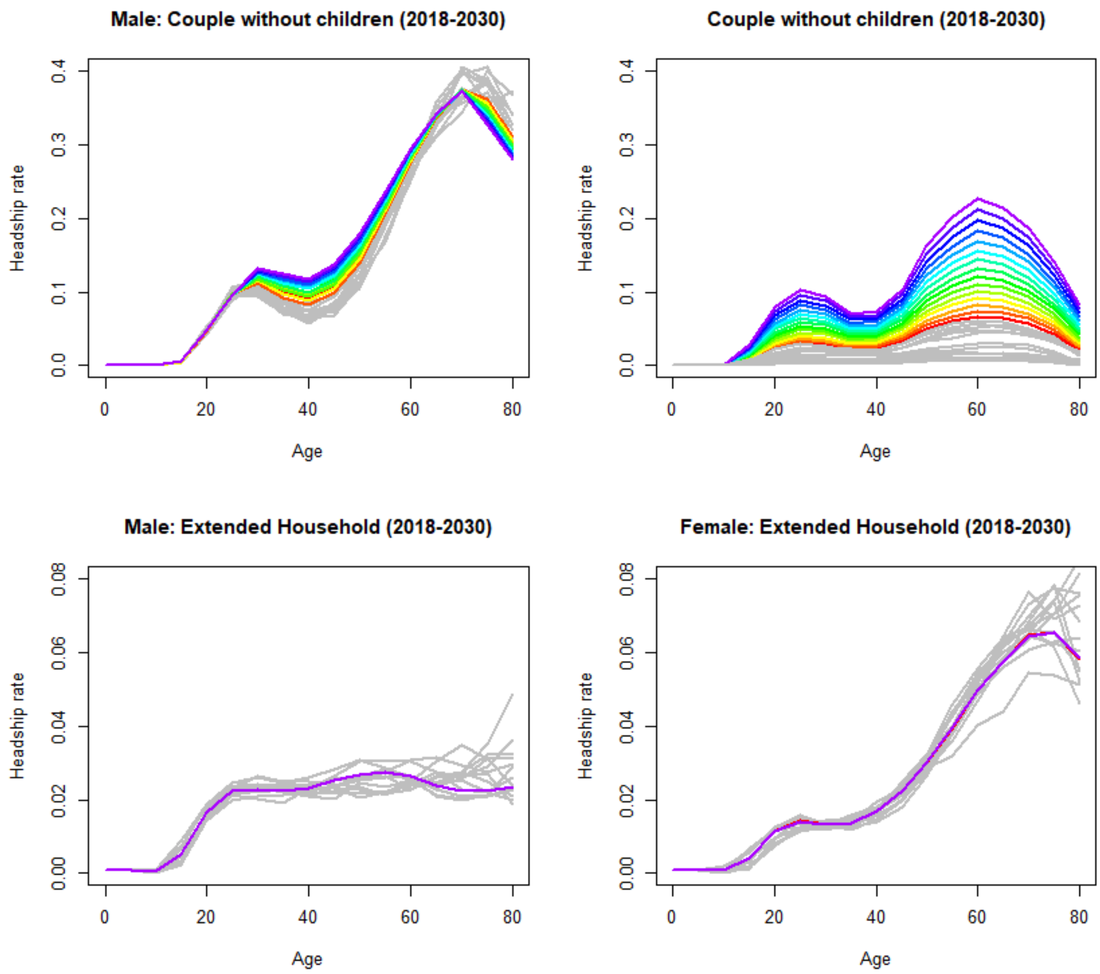


Source: Author's elaboration.

Similarly, Lee-Carter models were created for the 5 household types and for both sexes, resulting in 10 models. While it is not possible to identify a clear trend in (t), because the change over the past time were small, (t) is constant, resulting in the headship rate as the shape of ax (the mean). The forecast headship rates for all household types and sexes for the Brazilian South East are shown in Figure 23.

FIGURE 23 – Headship rate forecast by age, sex and household type, Brazilian South East Region (2018-2030)





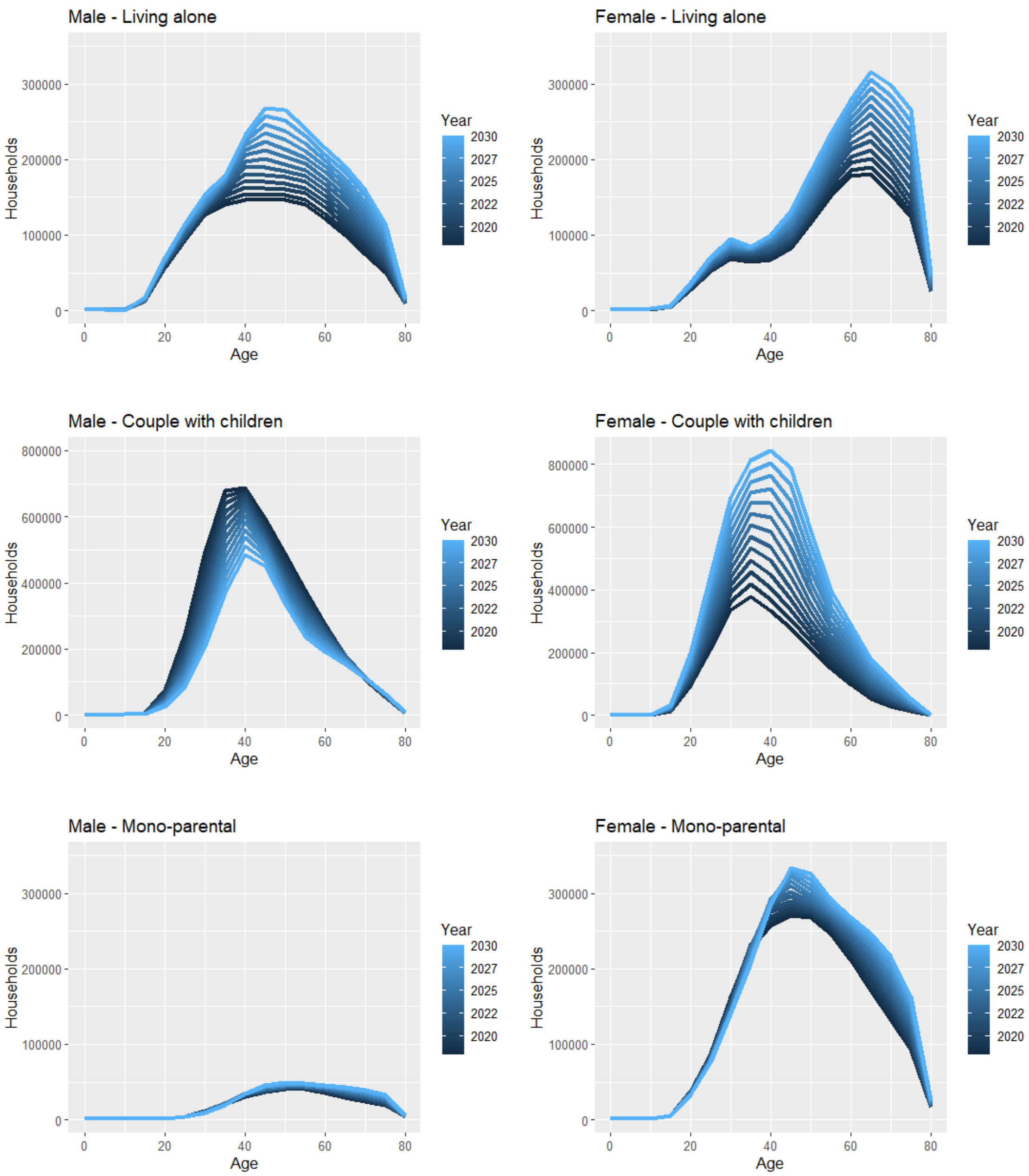
Source: Author's elaboration.

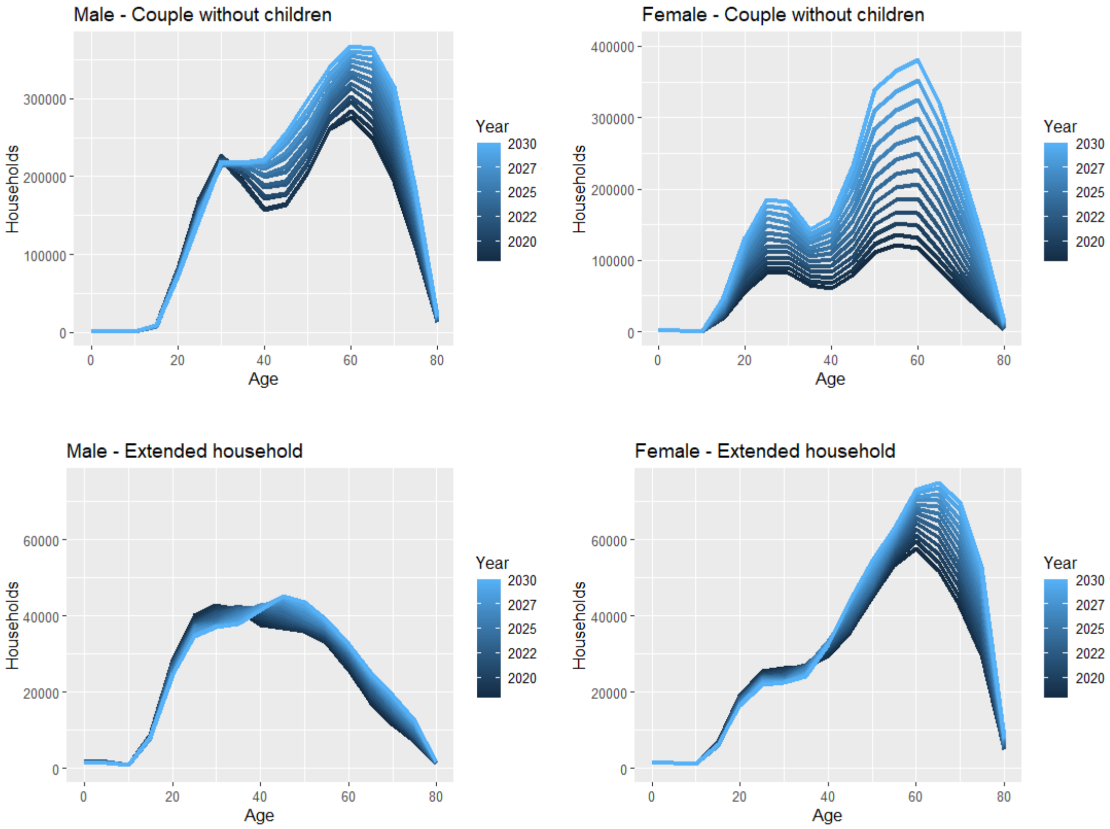
Once the headship rates for the future period and the population by Cohort Component were obtained, the multiplication yields the number of households for each type. The future number of households by household type, age and sex for the Brazilian South East are shown next.

3.5 Results: household projections for São Paulo State, by age, sex, and household type

This section is exclusively to present the result of the proposed household projection method. Discussed in Chapter 1, important areas of knowledge are interested in the number of households, but few projection methods present those results according to age, sex, and household type distributions. From Figure 24, it can be obtained exactly how many households is expected to increase or decrease in a specific age, of a specific sex and household type between two future years. This detailed information can derive a bunch of different application in social housing demand, family formation and environmental issues such as consumption of water, energy, and durable goods, and finally afford high focus public policies. Despite that, this information comes from a static projection method based in the classical Headship Rate Method, which means, it only requires cross sectional data sources and could be applied at any national or subnational Statistical Office for an official practice of household projection.

FIGURE 24 – Future number of households by sex and household type, São Paulo state, Brazil (2018-2030)





Source: Author's elaboration.

CHAPTER 4 – VALIDATION OF THE FPC HOUSEHOLD PROJECTION: HISTORICAL PROJECTION (2000-2010) AND 2010 CENSUS COMPARISON

As mentioned earlier, the technique of making historical projection is often used in the literature to validate projection models and it will be reproduced here. This Chapter will compare the results obtained by the proposed FPC method using past data (1980 and 1990 decades) to predict the 2010 census data. It simulates a user belonging to a statistical institution, a student or a consulting company member interested in making a household projection from 2000 to 2010, using only conventional and public Brazilian data sources from 1980 and 1990 decades, R open-source software and no further information on fertility, mortality, migration, and headship rates behavior beyond year 2000.

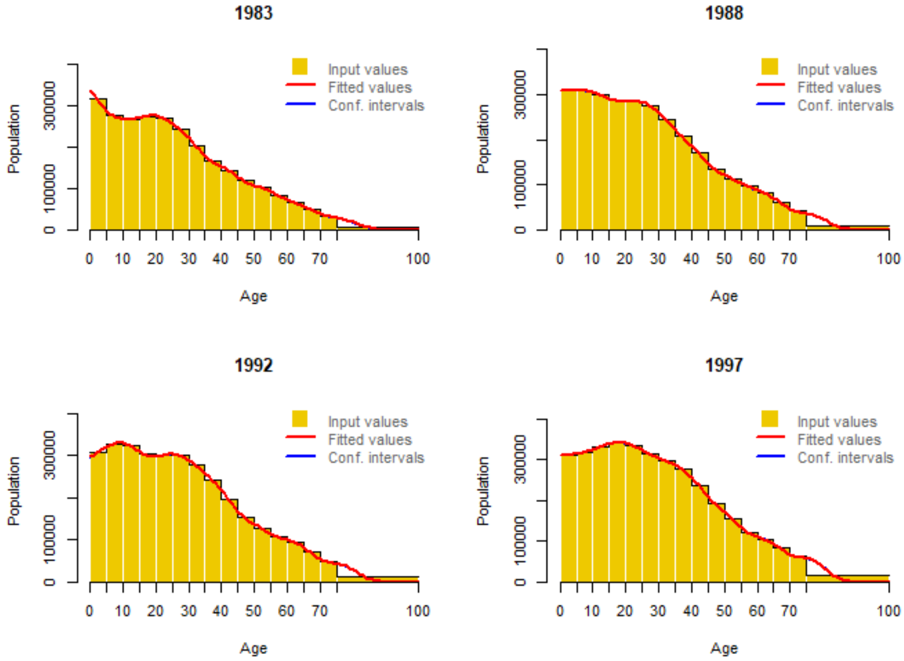
In addition to measuring errors from FPC method, this section will also highlight and discuss the main disadvantages and potentialities of this new approach. At the same time, suggestions will also be made for users interested in making population projections using Brazilian data.

The data required for FPC method application are the same as presented in Chapter 3: population data by single age for each observed year; the number of live births by single age of the mother for each observed year; the number of deaths and single age for each observed year; and the headship rate by single age and each observed year. For those who wish to reproduce this method using Brazilian data sources, this information is generally available in five-year ages group and not always for all years. Often the user must use additionally interpolation, disaggregation, and smoothing methods.

Particularly for this historical projection, the population data were obtained through SEADE Foundation for 1980 and 1990 decades, as the older time series are no longer available in the IMP/SEADE Foundation or DATASUS systems. Data on fertility and mortality from SINASC and SIM are only available for 1979 to 2019 and 1996 to 2019, respectively, on DATASUS website. So, the historical time series from 1980 to 2000 on fertility and mortality was also obtained under SEADE Foundation permission.

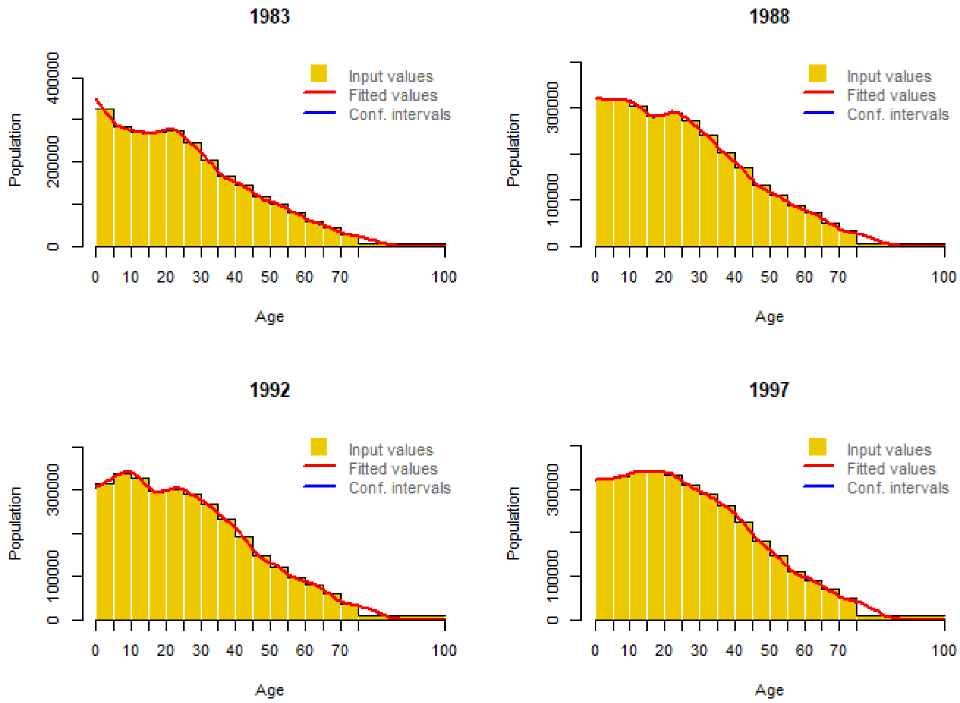
Figures 25 and 26 show the population results from disaggregating five-year ages groups to single ages obtained by SEADE Foundation using the Penalized Composite Link Model (RIZZI; GAMPE; EILERS, 2015; PASCARIU et al., 2018) present in the R package called “ungroup”. Years 1983, 1988, 1992 and 1997 were chosen as examples.

FIGURE 25 – Ungrouping 5-years age group population data to single age population data by Penalized Composite Link Model, Female, São Paulo, Brazil (1983, 1988, 1992 and 1997)



Source: Self-elaborated using Foundation SEADE data.

FIGURE 26 – Ungrouping 5-years population data to single age population data by Penalized Composite Link Model, Male, São Paulo state, Brazil (1983, 1988, 1992 and 1997)

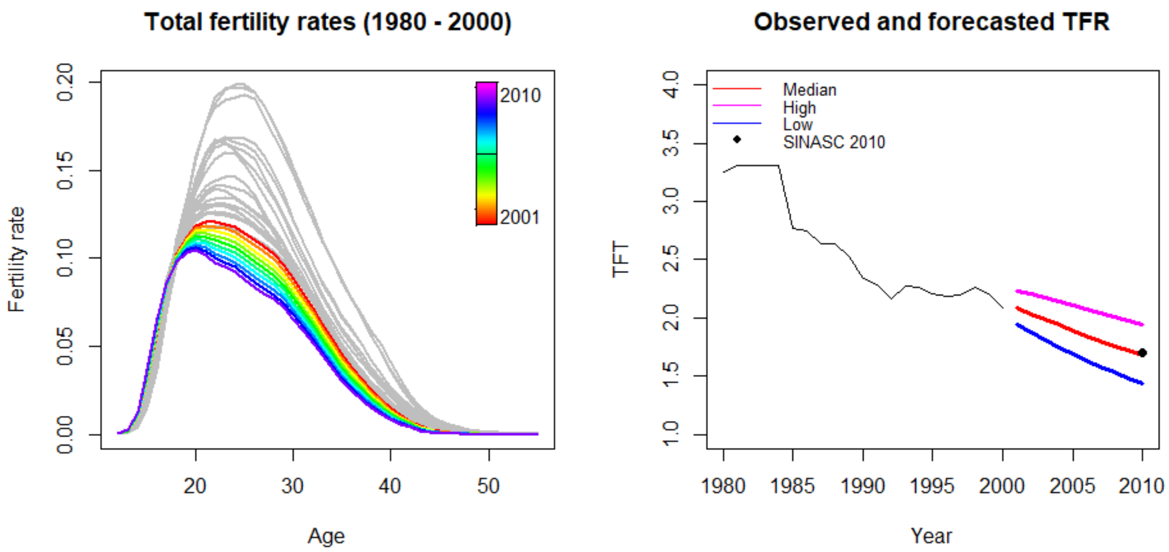


Source: Self-elaborated using Foundation SEADE data.

4.1 Comparing FPC historical fertility projection (2000-2010) and 2010 fertility

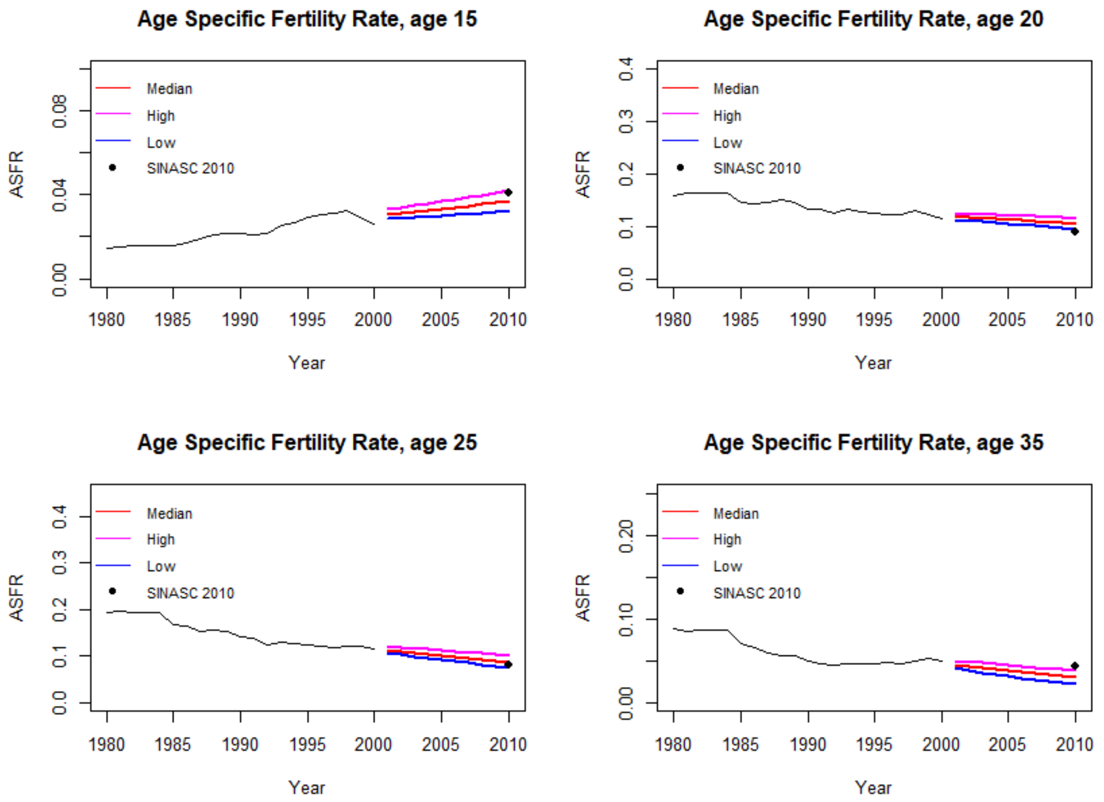
Following the same steps of section 3.1 for the fertility forecast, Figure 27 shows age-specific fertility rates and total fertility rate (TFR), observed (1980-2000 gray curves) and forecasted (colorful curves) for São Paulo state using FPC method. The black dot represents SINASC fertility estimation for 2010. Looking at age-specific fertility rates in graph 27, it is possible to affirm that FPC method has successfully forecasted the ASFR for most ages resulting a TFR equal to 1.68 with close proximity to TFR equal to 1.7 in 2010 SINASC data.

FIGURE 27 – Age-specific fertility rates and Total Fertility Rate (TFR), observed (1980-2000) and forecasted for São Paulo state, Brazil (2001-2010)



Source: Author’s elaboration using SINASC and Foundation SEADE data.

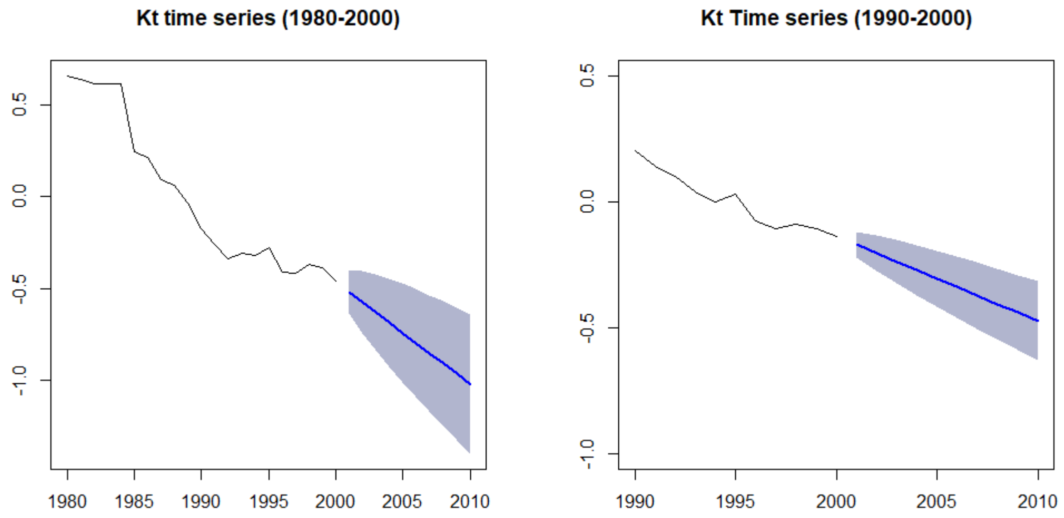
FIGURE 28 – Age-Specific Fertility Rates forecast by age 15, 20, 25 and 35. São Paulo state, Brazil (2001-2010)



Source: Author's elaboration using SINASC and Foundation SEADE data.

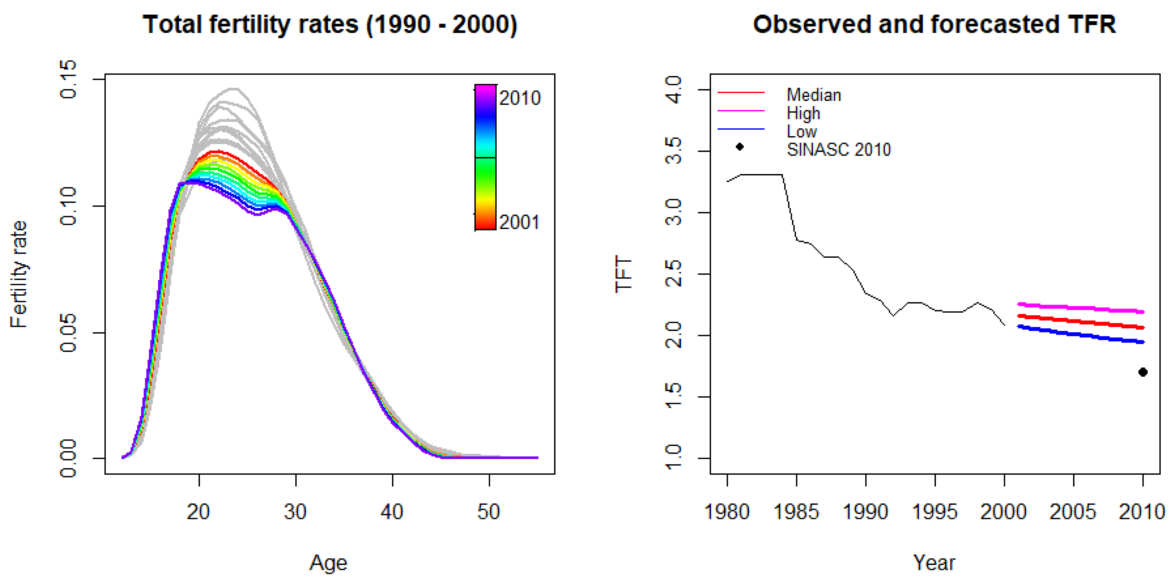
One important consideration about the method is that the fertility behavior over the observed period is not constant. It can vary from periods when the fertility drop is more intense to periods when fertility drop is less intense. Then, choosing different observed periods will produce different b_x and k_t effects, in other words, different forecasting. The Graph 28 compares the k_t effect using data from 1980-2000 and 1990-2000. The fertility decline over time is greater in mid-1980s, resulting in a higher inclination k_t compared to the fertility decline in 1990s. Thus, using only data from the 1990s would produce a lower fertility inclination trend overestimating the TFR of 2010. Figure 29 shows fertility forecasting using the k_t time series (1990-2000).

FIGURE 29 – Comparing k_t fertility trend using 1980-2000 and 1990-2000 time series, São Paulo state, Brazil



Source: Author's elaboration using SINASC and Foundation SEADE data.

FIGURE 30 – Age-specific fertility rates and Total Fertility Rate (TFR), observed (1990-2000) and forecasted for São Paulo, Brazil (2001-2010)

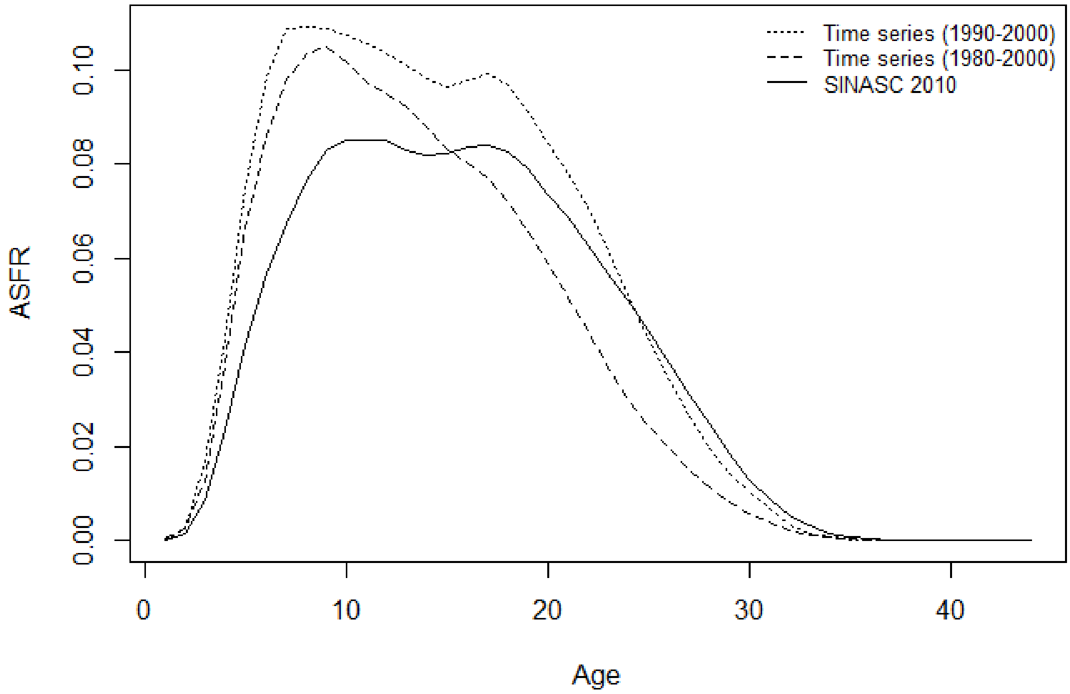


Source: Author's elaboration using SINASC and Foundation SEADE data.

Figure 31 compares the ASFR for 2010 using both time series 1980-2000 and 1990-2000 with SINASC 2010. When using the k_t time series from 1990-2000 we would have an age pattern much closer to that observed in SINASC 2010, showing that fertility age effect (b_x) was estimated more accurately than in the 1980-2000 time series, but at a higher level (TFR = 2.06) due to the smoother k_t.

Meanwhile, the forecast using k_t time series from 1980-2000 overestimate fertility at younger ages, since the b_x obtained from the 1980s considers a smaller drop in fertility at these ages compared to a time series using only the 1990s data. Besides that, the time effect (k_t) is better estimated using the 1980-2000 series reaching TFR equal to 1.69, closer to TFR 1.7 of SINASC. As the objective of this exercise is to simulate a user who does not know what happened in 2010, we stay with the complete time series (1980-2000).

FIGURE 31 – Comparison between SINASC age-specific fertility rate and forecast using 1990-2000 and 1990-2000 time series, São Paulo state, Brazil (2010)

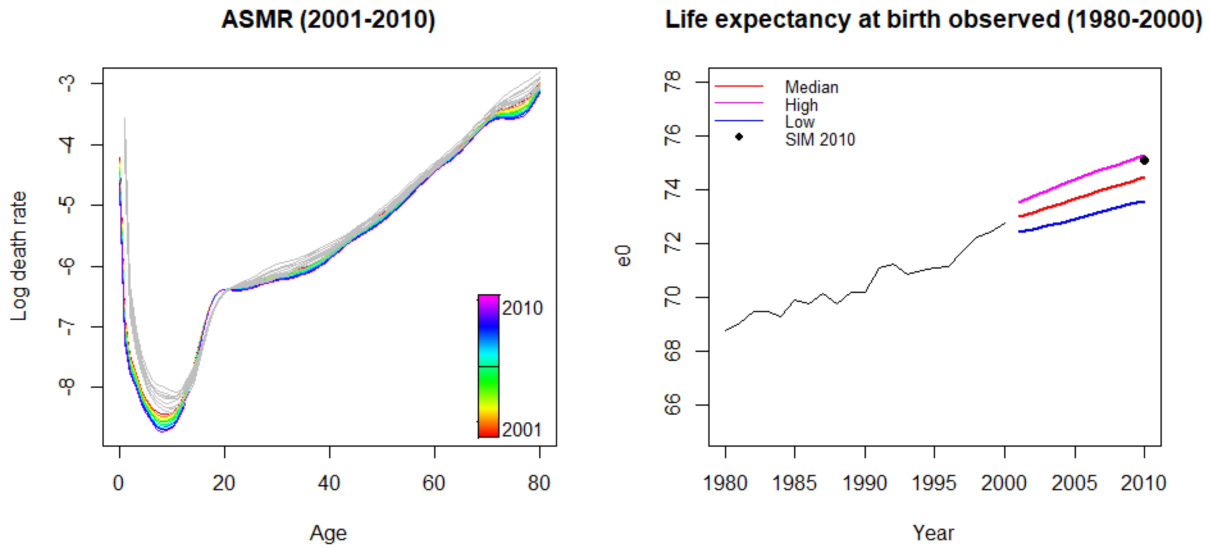


Source: Author's elaboration using SINASC and Foundation SEADE data.

4.2 Comparing FPC historical mortality projection (2000-2010) and 2010 mortality

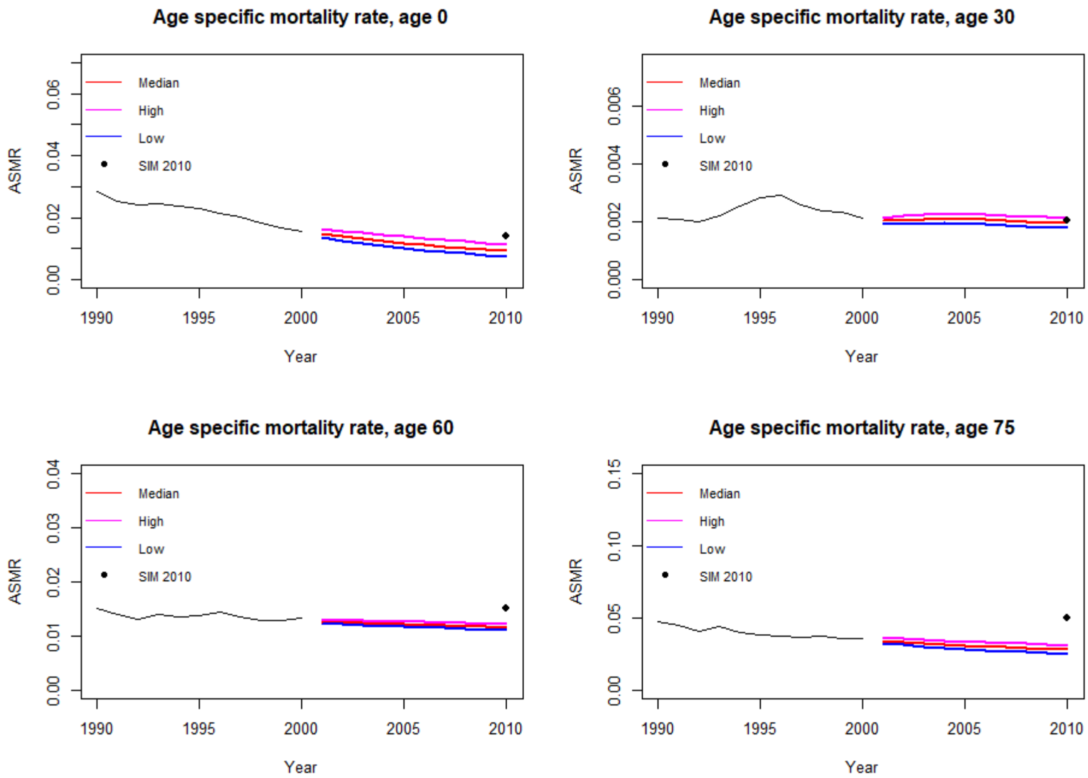
Following the same steps of section 3.2 for the mortality forecast, Figure 32 shows age-specific mortality rates and life expectancy at birth (e_0), observed (1980-2000 gray curves) and forecasted (colorful curves) for São Paulo state using FPC method. The black dot represents SIM mortality estimation for 2010. Looking at age-specific mortality rate in graph 32, it is possible to affirm that FPC method underestimated ASMR for most ages. The observed SIM life expectancy at birth is close to the higher confident interval scenario. Even so, it can be considered as a successful forecast, which the real value is within the confidence intervals.

FIGURE 32 – Age-specific mortality rates and life expectancy at birth, observed (1980-2000) and forecasted (2001-2010) for São Paulo, Brazil



Source: Author's elaboration using SIM and Foundation SEADE data.

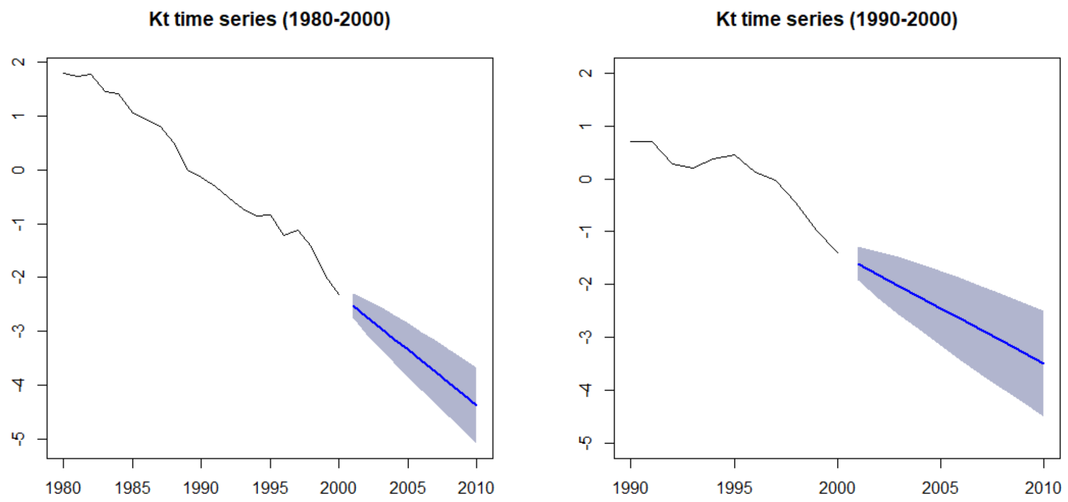
FIGURE 33 – Age-specific mortality rates forecast by age 0, 30, 60 and 75. São Paulo state, Brazil (2001-2010)



Source: Author's elaboration using SIM and Foundation SEADE data.

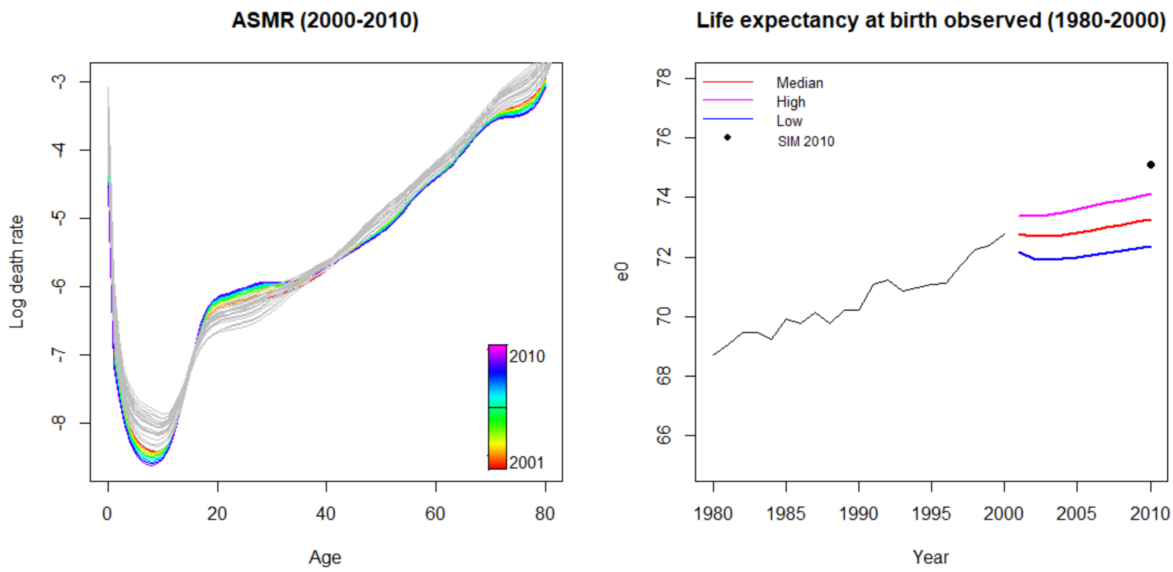
Comparing different time series, the k_t mortality effect is also smoother considering only the 1990-2000 period. Figure 34 shows that smoother mortality k_t estimate would provide higher mortality rates, consequently, lower life expectancy at birth compared to 2010 SIM value.

FIGURE 34 – Comparing k_t mortality trend using 1980-2000 and 1990-2000 time series, São Paulo state, Brazil



Source: Author's elaboration using SIM and Foundation SEADE data.

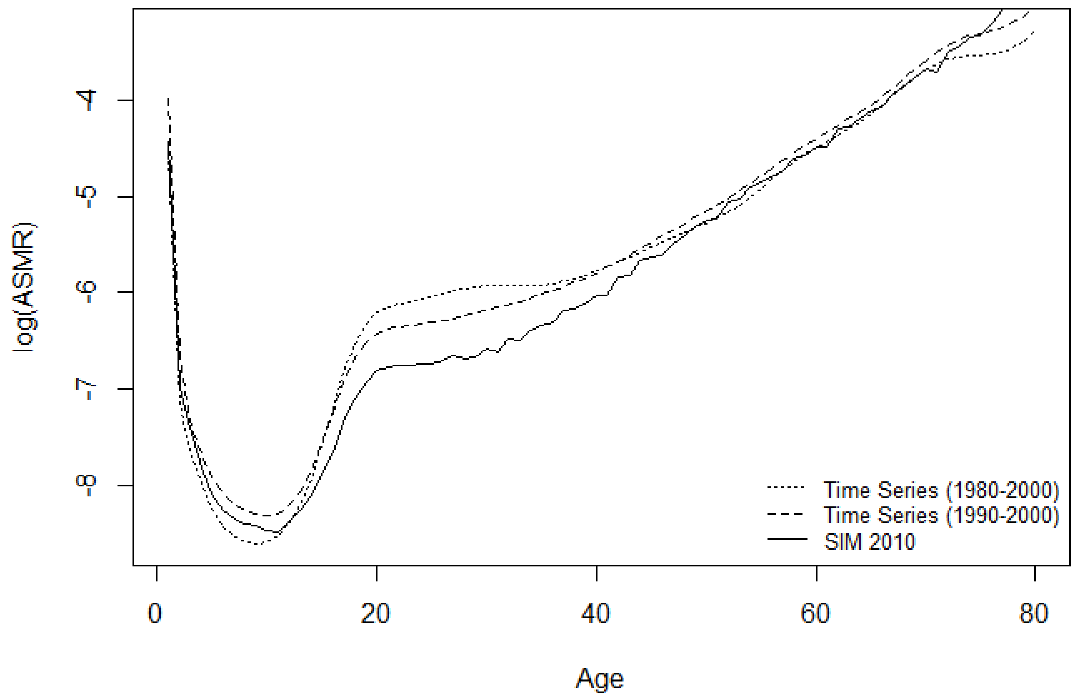
FIGURE 35 – Age-specific mortality rates forecast, São Paulo, Brazil (1980-2010)



Source: Author's elaboration using SIM and Foundation SEADE data.

Considering the 1980-2000 time series, the drop in mortality is greater, resulting in a higher life expectancy. However, this drop occurs at different ages (b_x) from the period 1990-2000. Therefore, if we considered the 1980-2000 time series for mortality, we would be underestimating life expectancy at birth as shown in Figure 36.

FIGURE 36 – Age-specific mortality rates forecast, São Paulo state, Brazil (1980-2010)

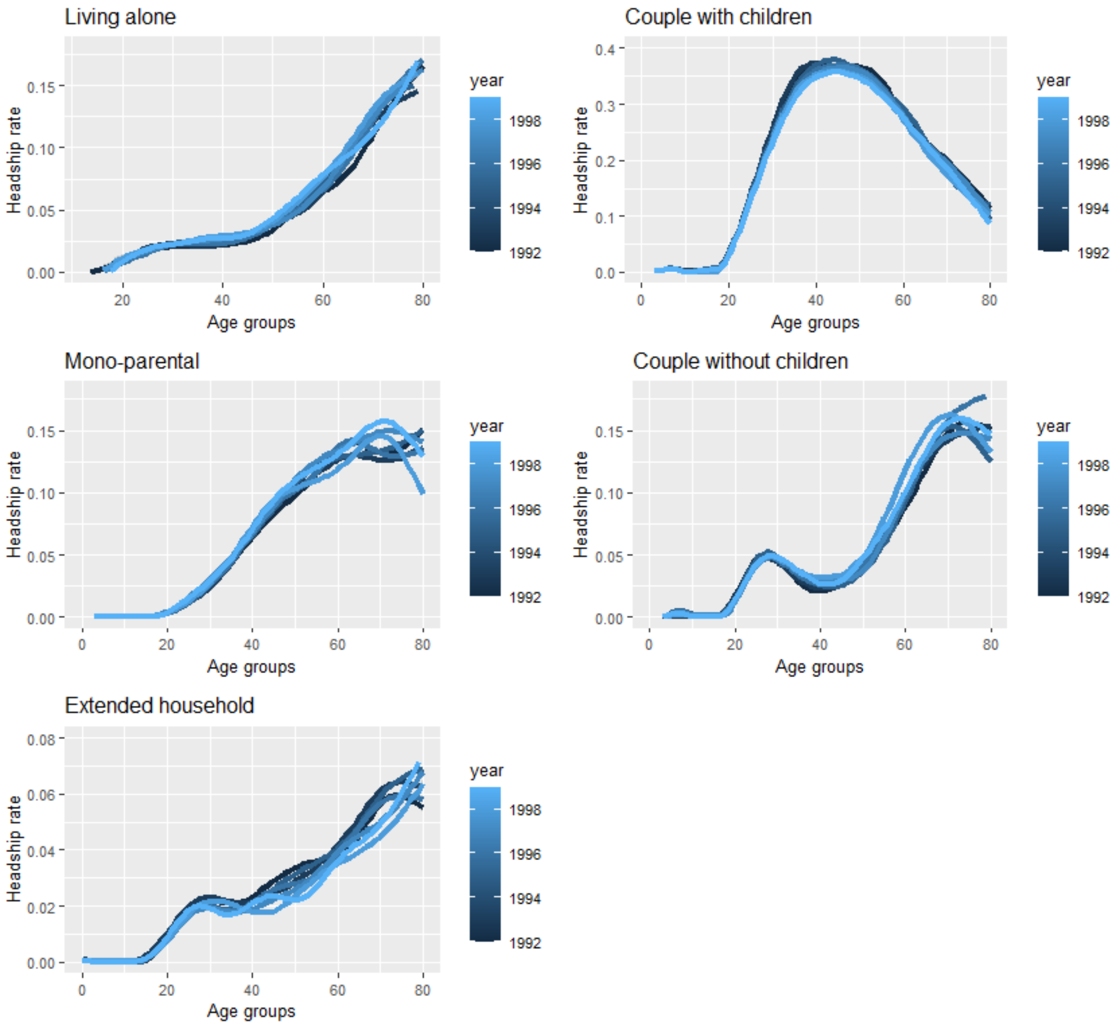


Source: Author's elaboration using SIM and Foundation SEADE data.

4.3 Comparing FPC historical headship rates projection (2000-2010) and 2010 headship rates

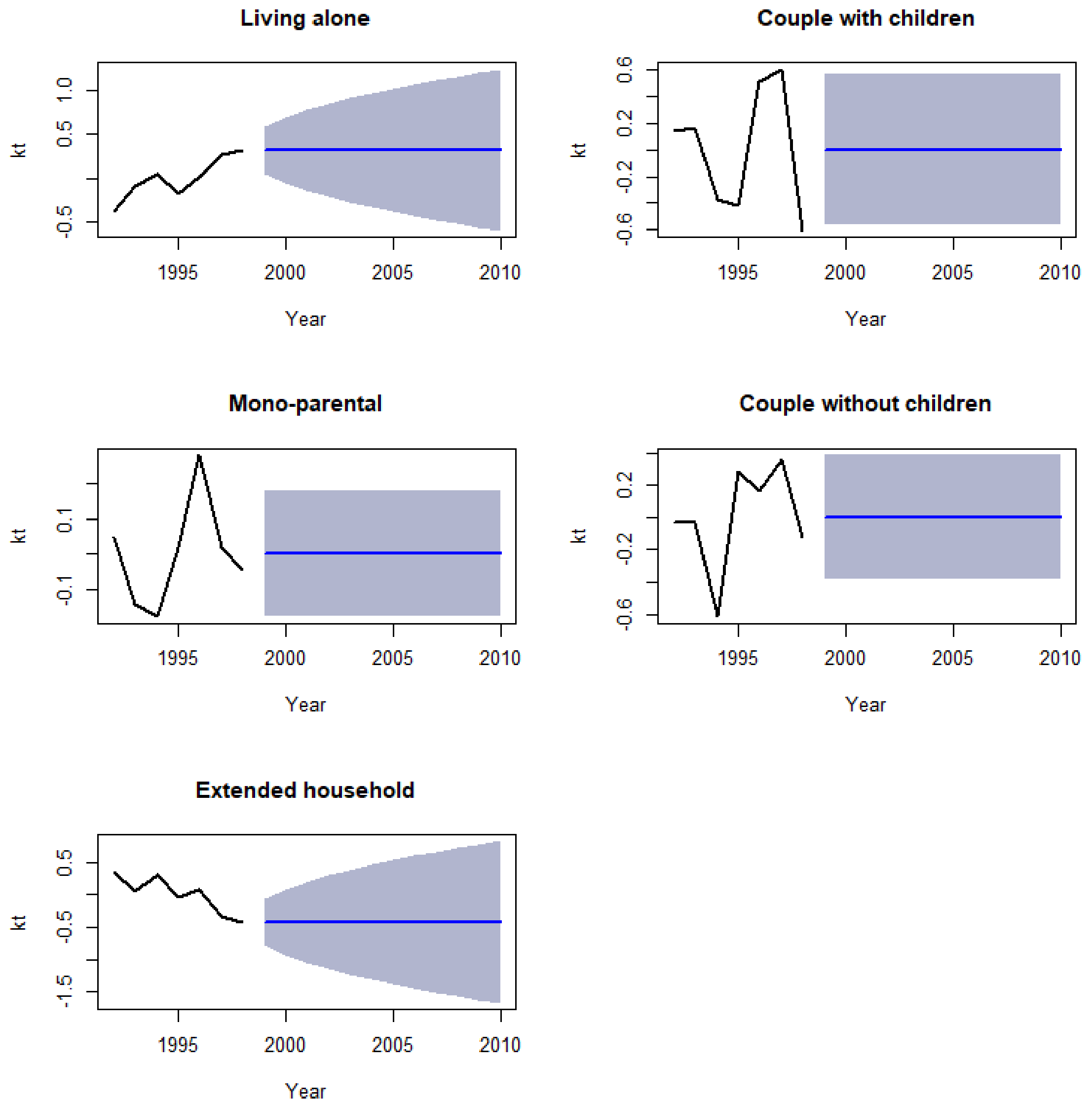
The user who is interested in forecasting headship rates in Brazil will face problems to find a data source for calculating headship rates over a long period of consecutive years. For the period of interest (1980-2000), as mentioned earlier, the PNAD's from 1992 to 1999 used the same survey question in the item "household reference person", therefore they have comparable results, which does not occur for the PNAD's from the 80's. Also, the 1994 PNAD was not carried out, consequently, we only have 7 points in time to obtain a trend in headship rates. Again, the population of the Brazilian Southeast was used because PNAD is a sample survey and the number of cases using only the state of São Paulo generates a great amount of variability in the data observation. Figure 37 shows the headship rates calculated using PNAD for the years 1992 to 1999 and Figure 38 shows the k_t trends using these 7 points in time.

FIGURE 37 – Headship rates by household type, Brazilian South East (1992-1999)



Source: Author's elaboration using PNAD (1992, 1993, 1995, 1996, 1997, 1998, 1999).

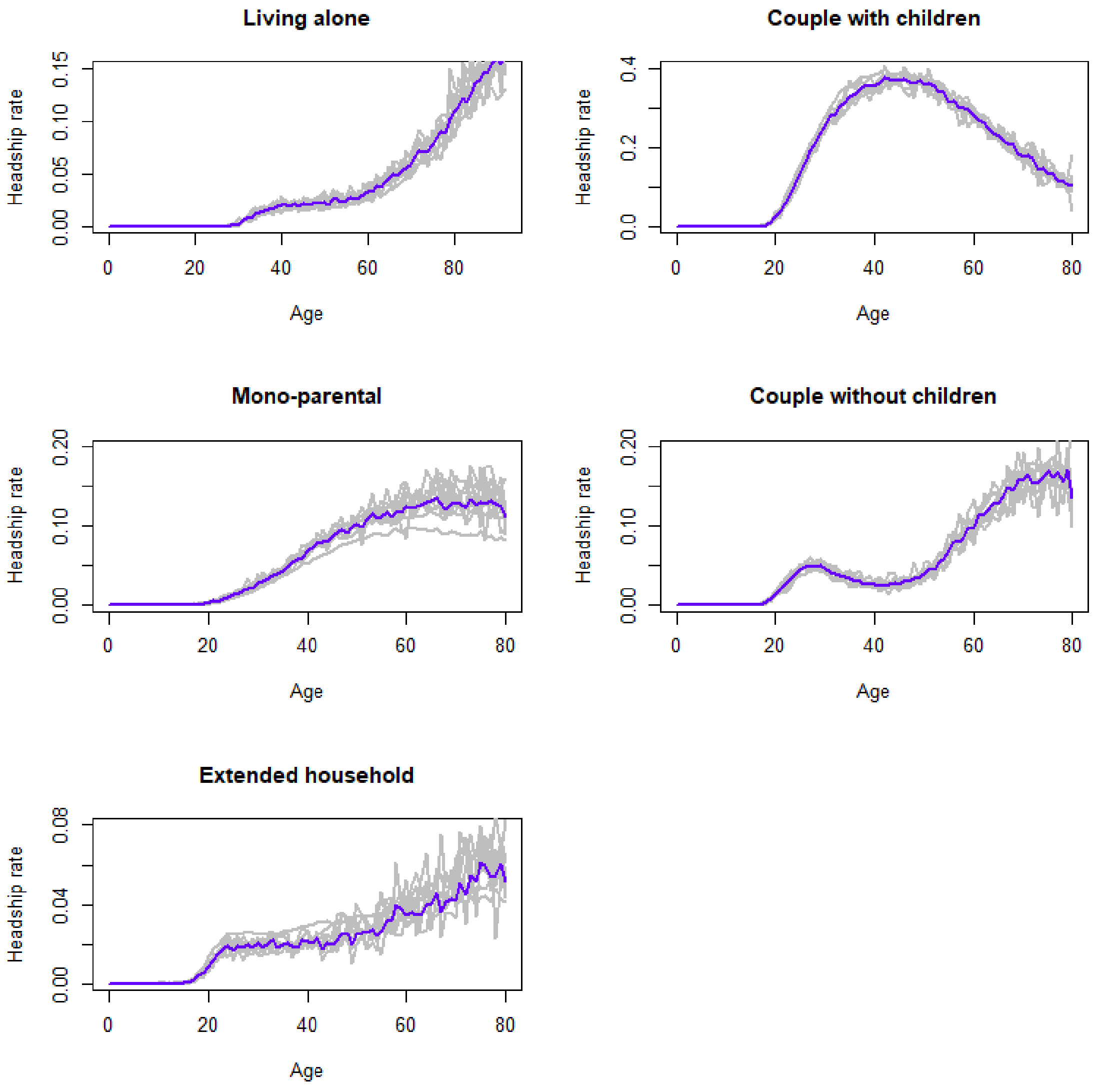
FIGURE 38 – Headship rate k_t trend by household type, Brazilian South East (1992-1999)



Source: Author's elaboration using PNAD (1992, 1993, 1995, 1996, 1997, 1998, 1999).

As we can see in Figure 37, it was not possible to establish a clear trend of headship rates change in any household type throughout the 90s decade. Two hypotheses may be occurring: most part of the headship rates changes may have occurred or have intensified after the 2000s. For example, the increase in living alone headship rates. The second hypothesis is the use of only 7 points in time was not enough to identify a clear pattern, due to the high variability. Facing this situation, the user would have to choose between keeping the k_t values constant, as in Figure 38, or choosing between a slight rate change within the confidence intervals. Figure 39 shows headship rates by household type if we chose to maintain the average headship rates of the 1990s.

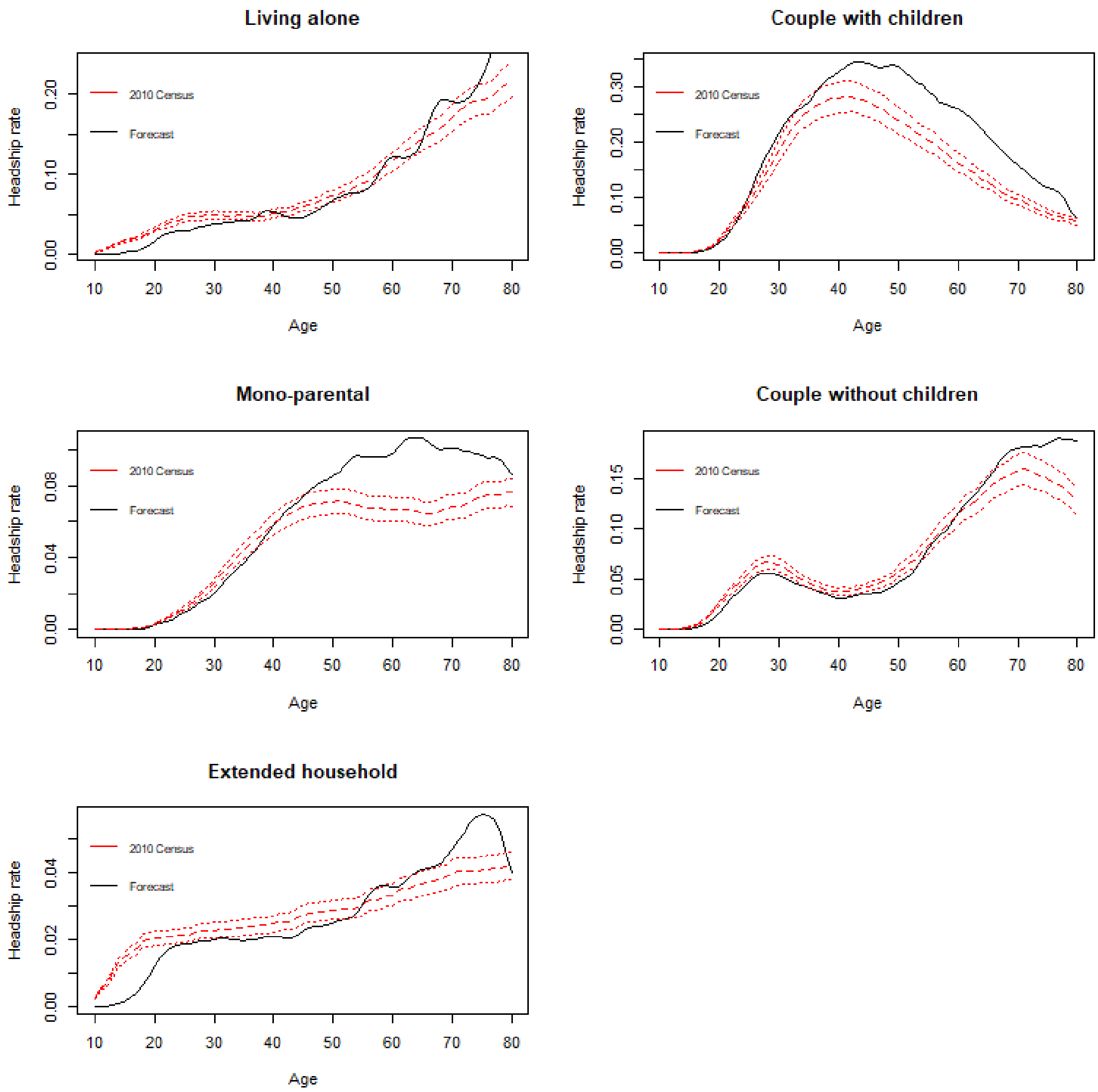
FIGURE 39 – Average headship rates by household type, Brazilian South East (1992-1999)



Source: Author's elaboration using PNAD (1992, 1993, 1995, 1996, 1997, 1998, 1999).

Comparing the average headship rates observed in the 90s with the headship rates observed in 2010, most of the error is concentrate in more advanced ages of households "couple with children", "mono-parental" and in younger ages of "living alone" and "extended household". In the following section, it will be demonstrated how these differences would impact the error in the total number of households.

FIGURE 40 – Headship rates comparison between 2010 census and forecasted, Brazilian South East

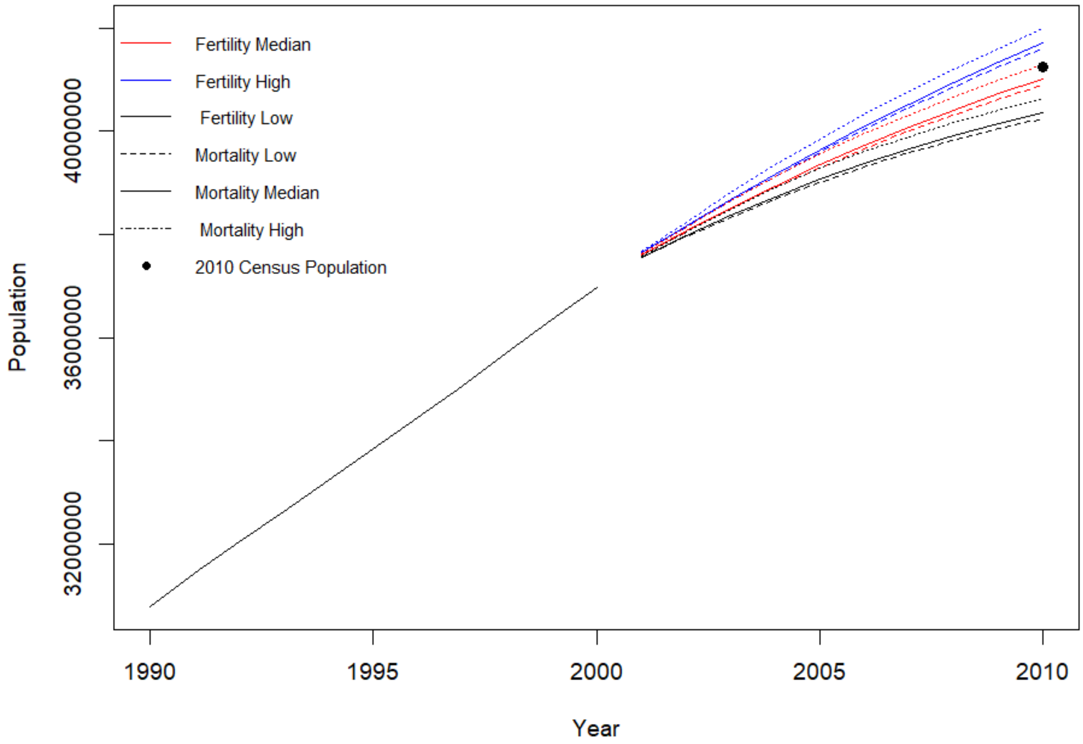


Source: Author’s elaboration using PNAD (1992, 1993, 1995, 1996, 1997, 1998, 1999 and 2010) Census.

4.4 Comparing FPC historical population projection (2000-2010) and 2010 census population

As soon as fertility, mortality and headship rates are forecasted by FPC method, the Cohort-Component Method can be made. Figure 41 shows the São Paulo total population observed data and all the variants of fertility and mortality scenarios base on 95% confidential intervals. The difference among 2010 census population (41,252,160) and medium fertility combined with medium mortality scenario (MedMed scenario = 41,008,410) is about 243,750, which represent a PE of 0.59%. It is under of 10% scenario mentioned by the literature (ZENG et al., 2014b; CAMPBELL, 2002; ERSI, 2007; KHAN; LUTZ, 2008).

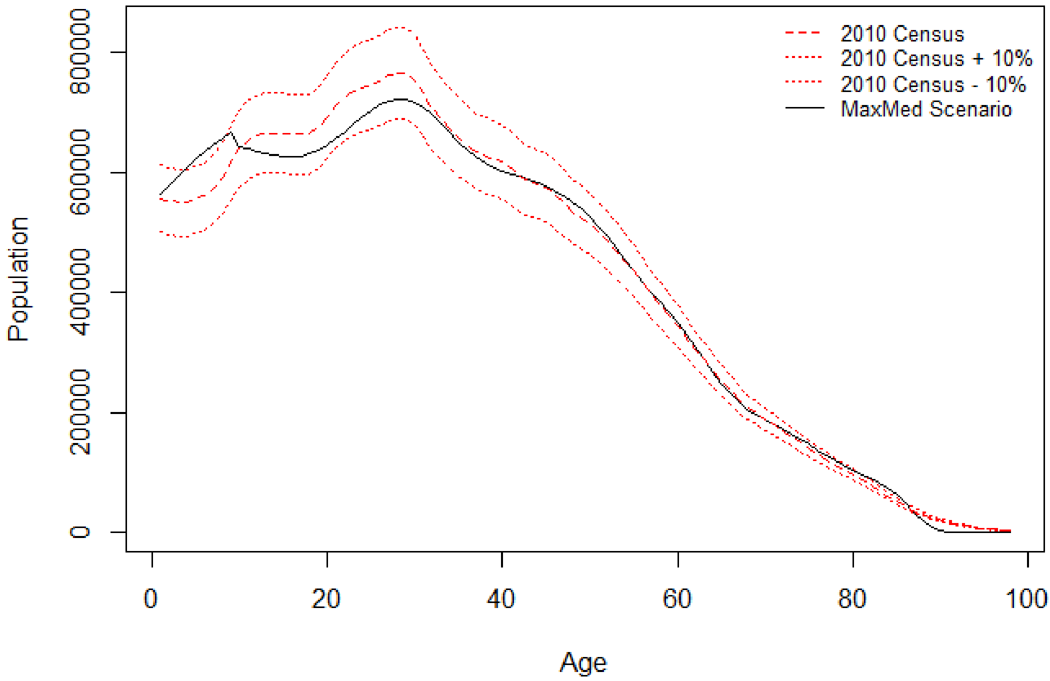
FIGURE 41 – São Paulo Total Population forecast by scenarios (2001-2010) and 2010 Census total population



Source: Author's elaboration.

Figure 42 shows the differences in age distribution. For most of the ages, the distribution forecasted by FPC method is under PE = 10% compared to 2010 Census age distribution.

FIGURE 42 – Age distribution forecast (1980-2010) and 2010 Census, São Paulo, Brazil

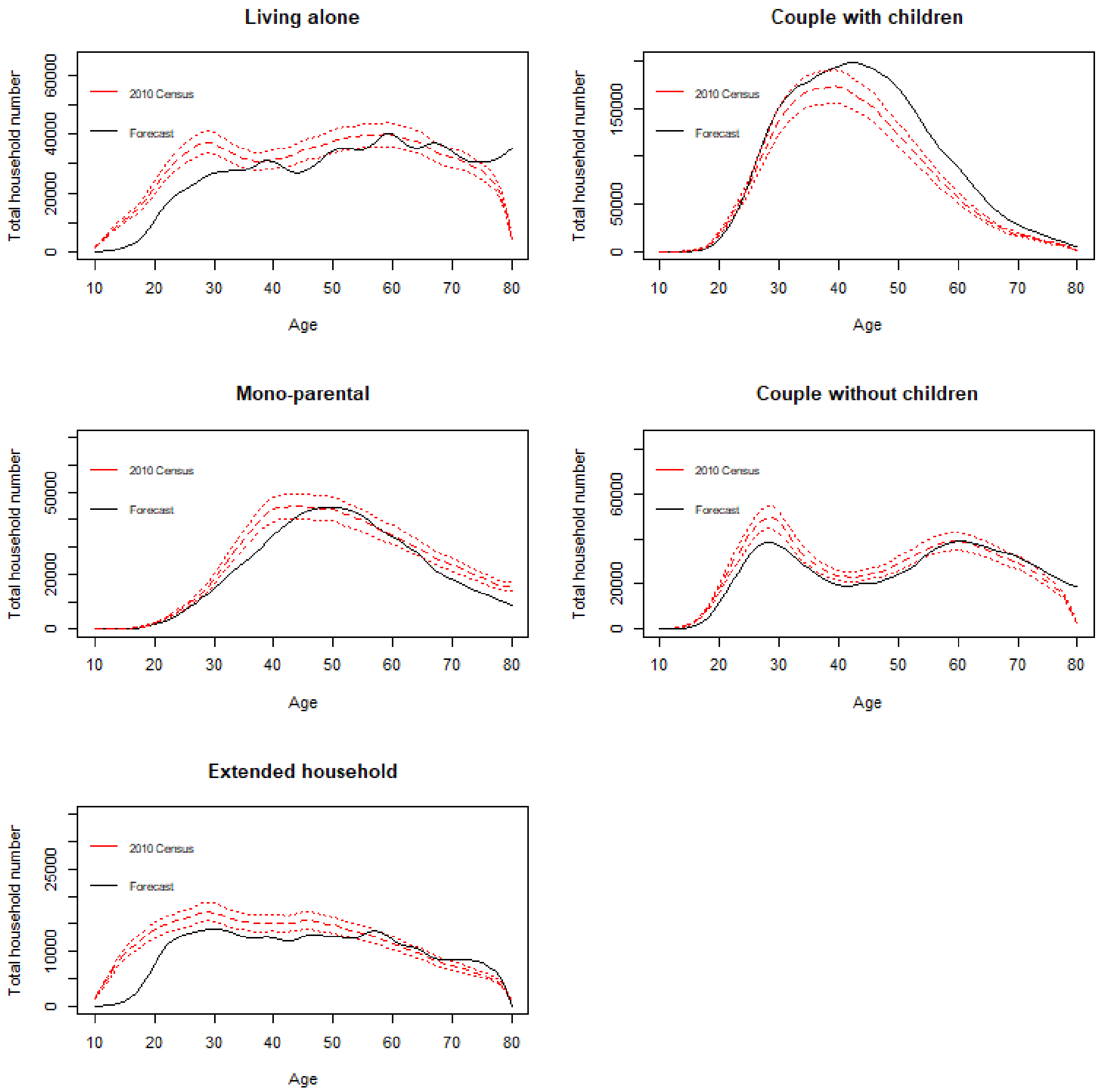


Source: Author's elaboration.

4.5 Comparing FPC historical household projection (2000-2010) and 2010 census household number

Finally, using the results from population projection and the average headship rates, it is possible to apply the future number of household heads formula from section 2.1. Figure 43 shows the comparison between future number of households by household type and age according FPC method and 2010 Census. Table 23 summarizes the total number of household's percentage error.

FIGURE 43 – Future number of households by type according FPC method and 2010 Census, São Paulo state, Brazil



Source: Author's elaboration.

TABLE 23 – Differences among total number of households* by FPC method and 2010 Census tabulation, São Paulo state, Brazil, 2010

Number of households by type	CENSUS	FPC	DIFF (#)	PE (%)
Living alone	2,118,258	1,874,169	244,089	11.52
Couple with children	5,267,670	6,465,828	1,198,157	11.97
Couple without children	1,917,852	1,749,691	168,161	8.77
Mono-parental	1,731,928	1,524,568	207,359	22.75
Extended household	837,920	697,152	140,768	16.80
Total	11,873,630	12,311,409	437,779	3.69

Source: Author's elaboration.
*Until 80 years old.
**Percentage error = [(Projection – 2010 Census) / 2010 Census] × 100.

4.6 Validation final remarks

The findings of this Chapter indicate that the proposed FPC method represents a valid way to project more detailed household information using headship rates with Brazilian data sources. The method produced low PE for the population and its age distribution and PE approximately 10% for total number of households by household type. The highest PE values occurs in total household number due to the strong assumption of constant headship rates since 2000.

In the same way to all projection methods, the quality of the projection depends on the quality of the input data. In this Lee-Carter-based method, data quality means having good estimates of annual rate curves during a sufficient period to identify a change pattern over time. In other words, the quality of the projection is associated with how close the ages effects (b_x) and time effects (k_t) are from the true values in the projection horizon. However, it is not always possible to obtain a considerable number of annual rate curves to estimate b_x and k_t satisfactorily.

Therefore, the great issue for a satisfactory estimation of b_x and k_t is the monotocity between the observed time series and the projection period. If the future period that we want to project has a similar decreasing pattern compared to the observed period, there are great chances of having a good estimate of b_x and k_t, consequently a good projection. On the other hand, sudden drops in ages effects (b_x) and time effects (k_t) will lead to unrealistic projection. We know that the trend observed in the historical period will not always be observed in the future. So, how different these time series are is the key for a good projection.

In the case of São Paulo state, the headship rates behavior captured by PNAD 1992-1999 was not satisfactory to predict the headship rates behavior from 2000 to 2010, considering them as constant. Part of this could be justified by the fact that the PNAD's sample size is too low (when smoothing the curves, important part of the information is lost due to the high variability of the data), partly because we have a short time series (only 7 annual headship rates observations before the decade 2000) and partly explained by more abrupt changes in the headship rates behavior happened after the 2000s. However, for users interested in making projections after 2000, Brazil has a longer historical PNAD's time series with the question about "who is the reference person in the household", providing more data to identify age and time effects (as how it was done in Chapter 3). Consequently, for recent projections the method tends to be more accurate.

In contrast to headship rates, the patterns of fertility and mortality rates have been identified with better success because SIM and SINASC datasets provide clear and constant patterns information on fertility and mortality rates throughout the 1980s and 1990s at different ages. In addition of having more points in time (every year since 1980) and larger sample sizes (as it is a Civil Register data), mortality and fertility rates have a monotonous decline during the period.

Another disadvantage of the method lies in the fact that the age effect remains constant throughout the projection (b_x). An extrapolation over such a long period using this method can lead to unrealistic age patterns. Some authors propose models that deal with the non-constancy of b_x that could be used to improve the performance of this method (LI; LEE; GERLAND, 2013).

CONCLUSIONS

The concern with choosing an appropriate household projection model for Brazil, expressed in this thesis, relies on the fact that the usage of household projections in Latin America differ from their usage in developing countries since it is necessary to consider intrinsic characteristics of the region. Firstly, the lack of national longitudinal research in the region restricts the range of possible methods to be applied. Second, Latin American data sources have greater enumeration errors, lack of coverage, and age heaping, as well as lack of continuity in historical time series, which adds a plenty of new assumptions and leads to greater projections errors. Third, some of the international household classification excludes the presence of households composed by non-relatives and other relatives, such as “extended households”, which its high percentage among Latin American countries is related to cultural intergenerational social support and survival strategies in times of crisis. Additionally, many demographic pace and shape patterns, such as leaving parent’s home, teenager fertility, young adult mortality, and consensual unions are significantly different from what is found in developed countries, the reason why it is not reasonable to replace this kind of information from other regions. Finally, Latin America has experienced rapid demographic changes in comparison to developing countries and has experiencing several economic and political crises, which can make times series difficult to predict using past data.

These are some of the reasons that could explain a huge difference in household projection practice between developing and developed countries showed in the literature review of Chapter 1. However, the comparison should not be limited only in terms of data and method used, but also what were the purposes behind each projection realized. In section 3 of Chapter 1, the literature review brought a reflection on “what is the household projection used for and for whom is the household projection used for?” It was argued that in Brazil, the fields that most demand household projection were some few academic studies and consultancies on electricity demand, social housing and housing market. The experiences were mostly restricted to the total number of households, not taking into account age and sex distribution, neither the size nor composition of these future households. This fact may be an indicator that household projection is being used for specific purposes and their methodological choices were not necessarily designed to match the demands of planning public policies, or necessarily reach a detailed projection method, or even provide open access to their results. For this reason, this thesis aimed to find and study household projection methods that could both be applied in Latin American context and provide more detailed information for different scientific and public purposes, such as, contribute to a regular and official household projection practice in national or sub-national statistic agencies.

In that sense, the accomplishment of three household projections based on conventional Brazilian data sources, two of them using static methods and one using a dynamic method, responds to the main objective of the thesis: Brazil has enough data source availability and data quality to support regular practice of household projection using the methods described here. Yet, the step toward a dynamic approach has not been fully validated, since the dynamic method applied, the ECCM, is a particular case of dynamic method in which it does not require data from longitudinal research. In addition, the thesis has an important role to report the differences between the methods based on empirical comparison results, while in general, studies in this field apply only one method and discuss why their method is better than others on theoretical base. It is important to mention that the objective of this thesis was not focused on discussing and analyzing the results of the future distribution of households in São Paulo state, but rather to show methodologically that there are possible ways for a practice of household projection in Brazil.

Both the Headship Rate Method (static) and the Extended Cohort-Component Method (dynamic) were able to provide complete information regarding the number, size, and households composition. They were validated in section 2.7 by a historical projection from 2000 to 2010 and compared to observed 2010 census data. The percentage difference between the ECCM and HRM projections and the 2010 Census observation for São Paulo state were less than 1% for the total population size, less than 2% for the total household number, approximately 0.05% for average household size and the projected age distribution stayed within the reasonable margin of PE 10%.

Using 5 different scenarios based on SEADE Foundation and IBGE’s projections, combining high/low mortality, high/low fertility, and constant rates, the projection towards 2050 pointed out a percentage change of 9.15% (HRM) and 11.05% (ECCM) in 2050 total population size compared to 2010 population size, while percentage change of 44.43% (HRM) and 44.96% (ECCM) in 2050 total number of households compared to 2010 total number of households. The results are consistent to the official SEADE Foundation projection and indicates that the total household number will still grow even in a low population growth context, due to the decreasing household size. In all 5 scenarios, smaller household types (living alone, couple without children and mono-parental) will be more frequent while couple with children and extended households tend to be less frequent.

The differences between ECCM and HRM in each scenario were less than 5%, however as Leiwen and O'Neill (2004) argue, the so-called “variant approach”, that is, the practice of varying a demographic component input to study the behavior of a projection result is difficult to reproduce in household projections and we cannot have a reasonable explanation for those differences between ECCM and HRM. According to the authors, in this type of projection there is a large number of input variables besides fertility, mortality and migration that are potential influencers of the results. Moreover, given the large set of results and different possible ways to investigate them, it is not clear which aspect of the results would be a reasonable parameter to define comparisons.

> It is not even clear which output variable should be used as the metric for the typical high/medium/low variants often produced in this approach. Should these be defined in terms of numbers of households, number of elderly households, age composition, multi-generation households, sex of household head, or what? Users have a range of different needs, and there is no single outcome that is of primary importance in all applications. Even if one were to select a single outcome as a basis for defining variants, the projection would be "probabilistically inconsistent" (LEE, 1999) in that the highest variant in terms of that outcome would not be the highest variant in terms of others. In summary, the weaknesses of the traditional variants approach in population forecasting (KEILMAN, 2003; LEE, 1999; ALHO, 1998) are compounded for household forecasts (LEIWEN; O’NEILL, 2004; p. 52).

Discussed in section 2.4, the Extended Cohort-Component Method (ECCM) displays an interesting tradeoff between having a more realistic model and having greater challenges in terms of data sources. The significant increased number of variables compared with the HMR adds a larger set of error sources, assumptions, and complexity. For example, it was not possible to directly apply a complete multistate life table with Brazilian Civil Register System because it has no record of widowhood, that is, there is no continuous record that counts the transition from married to widowed state, so the number of widowed persons is obtained through an indirect measure using information on mortality. In addition, Civil Register System concerns only marriages and divorces, it has no information about cohabiting-couples. Consequently, all cohabiting-couples were assumed to have the same transition rates of marriage couples. It is a rough approximation since Brazil and Latin America historically have high cohabiting proportions. For this reason, ProFamy’s module that consider 7 marital states, including all cohabitation status, cannot be used missing great potential of the method. The application of ECCM in other regions of Brazil seems to be even more challenging since Civil Register and Vital Statistics Systems still has important regional data quality differences.

Another example of the complexity of ECCM input for Brazilian data sources is when preparing inputs for base population. The categories defined by the software do not always correspond to the categories found in Brazilian censuses and some errors can be caused by the matching recodification. While the question about "relationship with the household head" had a set of 11 and 20 categories to classify the residents of the household, in the 2000 and 2010 Census, respectively, ProFamy requires only 8 categories. In the 2000 Census, there is no category named "grandfather" or "great-grandfather". So, those persons were included in the category "other relatives" in ProFamy's categorization, creating future problems in estimating 3 generation households.

Furthermore, discussed in section 2.7, although it is possible to derive more detailed household composition results, ECCM has their own "ego" marker and no additional information is available to reproduce ECCM ego's using census data. It makes harder to compare results with household type based on reference person and this can represent an important barrier to the widely usage of ECCM in official statistics offices.

Ultimately, the discussion among the possible methods applied for the Brazilian population on Chapters 1 and 2 leads to the main question "which method should an institution, consultor or researcher choose: a HRM or ECCM?" While ECCM enables the creation of more complex scenarios, more detailed results, allowing interesting sensitivity studies, but paying the price of more complex inputs, the HRM is simple, has few inputs, uses few data sources, has general results close to ECCM, although it has criticisms about not having a theoretical link between headship rates and demographic rates. The answer to this question is: there is not one method that is considered better than the other, but a set of assumptions, choices and objectives that may or may not be adequate. However, HRM will continue to be more frequently used because of its simplicity. Also, in the short and medium term, a longitudinal data source at the national level in Brazil still has remote chances to happen and we have been moving into the opposite direction of reducing the questionnaires sizes and massive spending cuts for census and statistical agencies funding. Therefore, Brazil will remain in the context of cross-sectional data sources and will hardly have better inputs availability for ECCM or any dynamic model.

From this perspective, the final part of the thesis intends to answer "how can we improve the classical HRM approach compared to the latest Brazilian household projection experiences?" It was found that many of these experiences have the strong and unlikely assumption that the headship rates will remain constant over time, they present outdated methods of mortality and fertility projections or use old population projections. Also, they often use private software that is difficult to access, reproduce and disseminate the results, they do not deal with probabilistic uncertainty, generally their scenarios are based on intuitions (even if based on past data) and they do not respond to the main demands on the number, size, and composition distribution by age and sex.

In Chapter 3, a new alternative method was proposed to improve these points. This method, called Functional Principal Components (FPC), is based on Hyndman and Ullah (2007) and Lee-Carter method (1992) for mortality and fertility forecast, but here, it is also adjusted for headship rate forecasting. It starts from PNAD's headship rates, to avoid problems with the Portuguese term used to identify the household head, and it derives an age effect (b_x) and a time effect (k_t) to identify a headship rate trend by household type, age, and sex. It can avoid the usual assumption of constant rates. The mortality, fertility and headship rates scenarios were created based on 95% confidence intervals and it is all programmed in software R, to avoid the usage of private software and the user can control the entire projection process, such as manipulate the dataset or plot their own results.

In Chapter 4, the FPC method was validated by a historical projection from 2000 to 2010 compared to 2010 census and it produced low PE for the population and its age distribution and PE approximately 10% for total number of households by household type. Mortality and fertility estimation from SIM and SINASC also were within the 95% confidence intervals. However, the 1992-1999 headship rate time series from PNAD was not satisfactory to predict the headship rates behavior from 2000 to 2010, considering all of them as constant during the projection period. Part of this could be justified by the fact that the PNAD's sample size is too low (when smoothing the curves, important part of the information is lost due to the high variability of the data), partly because we have a short time series (only 7 annual headship rates observations before the decade 2000) and partly explained by more abrupt changes in the headship rates behavior happened after the 2000s. Even though, the final total number of households by type and age distribution were close to the observed in 2010 census. The main disadvantages of the FPC method are that it relies on monotonicity between the observed time series and the projection period, the age effect remains constant throughout the projection and it is not a coherent forecast, i,e, each household type projection is independent, not necessarily matching when it summed.

As a final suggestion to an official statistic office which cannot afford a longitudinal research and intent to improve their surveys towards dynamic household projections, they should apply questions about the time the individuals are in each marital state, including consensual unions. This question can be made by asking the age of the person when starting the current marital state, which already had been used in Brazilian 1991 census, or what was the marital state one year before the census reference date. With this information we could estimate the transition rates between marital status and consensual unions even in a cross-sections perspective. Not only for dynamic household projection but it also would be useful for other demographic tools such as event history analysis and multi-state life Tables.

Finally, the year of 2020 and 2021 was marked by a pandemic caused by COVID-19, which will have an important impact in every demographic component: mortality, fertility, migration as well as household patterns. The population and household projection models must take it into account, and they are also an important tool to understand new scenarios ahead.

REFERENCES

ABRAINC – ASSOCIAÇÃO BRASILEIRA DE INCORPORADORAS IMOBILIÁRIAS; FUNDAÇÃO GETÚLIO VARGAS. **Análise das necessidades habitacionais e suas tendências para os próximos dez anos.** São Paulo, SP, 2018. (Produto 2 – Relatório Técnico Final – 2ª Versão).

ABURTO, J. M.; RIFFE, T.; CANUDAS-ROMO, V. Trends in avoidable mortality over the life course in Mexico, 1990–2015: a cross-sectional demographic analysis. **BMJ Open,** London, v. 8, n. 7, p. e022350, 2018.

AIDAR, T. et al. Age of children leaving home in São Paulo State, Brazil: tendencies, sex and race differences from 1991 to 2010. In: INTERNATIONAL POPULATION CONFERENCE – IUSSP, 28., Cape Town, South Africa. **Anais…** França: IUSSP, 2017.

ALHO, J. M. **A stochastic forecast of the population of Finland.** Helsinki: Statistics F., 1998.

ALKEMA, L. et al. Probabilistic projections of the total fertility rate for all countries. **Demography,** New York, NY, v. 48, n. 3, 815-839, 2011.

ANDRADE, F. V.; PINHEIRO, R. B. Análise de decomposição da Projeção de Consumo de Energia Elétrica no Brasil para o setor residencial. **ENGEVISTA,** Niterói, RJ, V. 16, n. 4, p. 340-355, 2014.

APARICIO, C. A. P.; BRUSSE, G. P. L. Demanda demográfica de viviendas: proyección de los arreglos residenciales hasta 2030 a partir de la población destinataria de un programa de vivienda social de la Compañía de Desarrollo Habitacional y Urbano (CDHU) en el estado de São Paulo. **Notas de Población,** Santiago de Chile, Chile, n. 107, p. 103-129, 2018.

AUGUSTO, M. F. L. Tábuas de nupcialidade legal para o Brasil–Estados da Região Sul. In: ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS, 14., 2004, Caxambu, MG. **Anais...** Belo Horizonte, MG: ABEP, 2004.

BAUDISCH, A.; VAUPEL, J. W. Getting to the root of aging. **Science,** Washington, DC, v. 338, n. 6107, p. 618-619, 2012. 338, pp. 618-619. (This short article summarizes why Hamilton was wrong: senescence is not inevitable).

BAY, G. Reflexiones sobre las estimaciones y proyecciones de población en América Latina. In: CAVENAGHI, S. (org.). **Estimaciones y proyecciones de población en América Latina:** desafíos de una agenda pendiente. Rio de Janeiro, RJ: ALAP, 2012. p. 51-86. (Serie e-Investigaciones, n. 2).

BENGTSON, V. L.; ROSENTHAL, C.; BURTON, L. Families and aging: diversity and heterogeneity. In: BINSTOCK, L. K. et al. **Handbook of aging and the social sciences.** Cambridge: Academic Press, v. 3, 1990. p. 263-287.

BENNET, J.; DIXON, M. **Single person households and social policy: looking forward.** New York, NY: Joseph Rowntree Foundation, 2006.

BENNETT, N. G.; HORIUCHI, S. Estimating the completeness of death registration in a closed population. **Population Index,** Princeton, NJ, v. 47, n. 2, p. 207-221, 1981.

BERGERON-BOUCHER, M. P. et al. Coherent forecasts of mortality with compositional data analysis. **Demographic Research,** Germany, v. 37, n. 17, p. 527-566, 2017.

BERMÚDEZ PARRADO, S.; HERNÁNDEZ RODRÍGUEZ, J. A.; PLANELLES ROMERO, J. **Proyecciones derivadas:** proyección de los hogares y proyección de la población activa de Andalucía. Sevilla, Espanha: Instituto de Estadística y Cartografía de Andalucía, 2011.

BERQUO, E. Reprodução após os 30 anos no estado de São Paulo. **Novos Estudos CEBRAP,** São Paulo, SP, n. 100, p. 9-25, 2014. Available from: https://www.scielo.br/j/nec/a/F5MYqd6p4Gm9NJ7FgJBX3Vh/?lang=pt.

BERQUÓ, E.; XENOS, P. **Family systems and cultural change.** Oxford: Clarendon Press, 1992.

BILAC, E. D. Transformações nos arranjos domiciliares no Brasil: interpretações possíveis. In: Berquó, E. (org.). **Demografia na Unicamp:** um olhar sobre a produção do Nepo. Campinas, SP: Editora da Unicamp, 2017.

BONGAARTS, J. The projection of family composition over the life course with family status life Tables. In: BONGAARTS, J.; BURCH, T. K.; WACHTER, K. (ed.). **Family demography:** methods and applications. Oxford: Clarendon Press, 1987. p. 189-212.

BONGAARTS, J.; BURCH, T. K.; WACHTER, K. W. (ed.). **Family demography:** methods and their application. Oxford: Clarendon Press, 1987. (International Studies in Demography).

BONGAARTS, J. The formal demography of families and households: na overview of the proximate determinants. **ISSUP Newsletter,** London, n. 17, p. 27-42, 1983.

BOOTH, H.; HYNDMAN, R. J.; TICKLE, L. Prospective life Tables. In: CHARPENTIER, A. (ed.). **Computational actuarial science with R.** [*S. l.*]: Chapman and Hall/CRC, 2014. p. 319-344.

BRASIL. Ministério de Minas e Energia **Projeção da demanda de energia elétrica para os próximos 10 anos (2017-2026).** Brasília, DF, 2017. (Nota Técnica DEA 001/17/Série Estudos da Demanda).

BRASIL. Ministério de Minas e Energia. Empresa de Pesquisa Energética. **Plano Nacional de Energia 2030.** Brasília, DF, 2007.

BRASS, W. The formal demography of the family: an overview of the proximate determinants. In: BRITISH SOCIETY FOR POPULATION STUDIES. **The family.** London: Office of Population Censuses and Surveys, 1983. p. 37-49. (OPCS Occasional Paper, 31).

BRASS, W. **Methods for estimating fertility and mortality from limited and defective data.** Chapel Hill: The University of North Carolina, International Program of Laboratories for Population Statistics, 1975.

BRASS, W. **Métodos para estimar la fecundidad y la mortalidad en poblaciones con datos limitados.** Santiago de Chile, Chile: Centro Latinoamericano de Demografia, 1974.

BURCH, T. K. Computer modeling of theory: Explanation for the 21st century. **PSC Discussion Papers Series,** London, v. 13, n. 4, 1999.

BURCH, T. K. Household and family demography: a bibliographic essay. **Population Index,** Princeton, NJ, v. 45, n. 2, p. 173-195, 1979.

CAMARANO, A. A.; KANSO, S. Como as famílias brasileiras estão lidando com os idosos que demandam cuidados e quais as perspectivas futuras? A visão mostrada pelas PNADs. In: CAMARANO, A. A. (ed.). **Cuidados de longa duração para a população idosa:** um novo risco social a ser assumido? Rio de Janeiro, RJ: IPEA, 2010. p. 93-122.

CAMPBELL, P. R. **Evaluating forecast error in state population projections using census 2000 counts.** Washington, DC: U.S. Census Bureau, 2002. (Population Division Working Paper Series, n. 57).

CANNAN, E. The probability of a cessation of the growth of population in England and Wales During the next century. **The Economic Journal,** Cambridge, v. 5, n. 20, p. 505-15, 1895.

CASTILLO, A. E.; PEÑA, L. S.; GUARDIÁN, G. P. Hogares y energía eléctrica en México. **Revista Espinhaço,** Diamantina, MG, v. 5, n. 2, p. 30-43, 2016.

CAVENAGHI, S.; ALVES, J. E. D. **Mulheres chefes de família no Brasil:** avanços e desafios. Rio de Janeiro, RJ: ENS-CPES, 2018. (Estudos sobre Seguro, n. 32).

CAVENAGHI, S.; ALVES, J. E. D. Domicilios y familias en la experiencia censal del Brasil: cambios y propuesta para identificar arreglos familiares. **Notas de Población,** Santiago de Chile, Chile, n. 92, 2011.

CEPDE – CENTRO DE ESTUDIOS DE POBLACIÓN Y DESARROLLO. **Proyecciones de los hogares cubanos 2015 – 2030**. Córdoba, 2016.

CHRISTIANSEN, S. G; KEILMAN, N. Probabilistic household forecasts based on register data-the case of Denmark and Finland. **Demographic Research,** Germany, v. 28, n. 43, p. 1263-1302, 2013.

CGIAE – COORDENAÇÃO GERAL DE INFORMAÇÕES E ANÁLISE EPIDEMIOLÓGICA. **Consolidação da base de dados de 2011:** Sistema de Informações sobre Mortalidade – SIM. Brasília, DF: CGIAE/SVS/MS, 2011.

CELADE – CENTRO LATINOAMERICANO DE DEMOGRAFIA. **Métodos para proyecciones demográficas.** San José, Costa Rica. 1984.

CEDEPLAR – CENTRO DE DESENVOLVIMENTO E PLANEJAMENTO REGIONAL. **Projeção do total de domicílios, por idade e sexo, para o Brasil e Unidades da Federação, 2003-2023.** Belo Horizonte,MG, 2007.

COALE A. J.; DEMENY, P.; VAUGHAN, B. **Regional model life Tables and stable populations.** New York, NY: Academic Press, 1983.

COALE, A. J.; TRUSSELL, T. J. Model fertility schedules: variations in the age structure of childbearing in human populations. **Population Index,** Princeton, NJ, v. 40, n. 2, p. 185-258, 1974.

COLCHERO, F. et al. The emergence of longevous populations. **Proceedings of the National Academy of Sciences,** Washington, DC, v. 113, n. 48, p. E7681-E7690, 2016.

CORTEZ, B. F.; LAZO, A. V.; MAGALHÃES, M. S. A nupcialidade legal no Brasil e nas Grandes Regiões: uma análise utilizando tábuas de vida de múltiplos-estados. In: ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS, 16., 2008, Caxambu, MG. **Anais...** Belo Horizonte, MG: ABEP, 2008.

CORTEZ, B. F. **A nupcialidade como um processo dinâmico:** uma abordagem markoviana. 2007. 125f. Dissertação (Mestrado) – ENCE/IBGE, Rio de Janeiro, RJ, 2007.

CRAICE, C.; GUERCIO, F. M. F. D.; D’ANTONA, A. O. A influência de características do domicílio no consumo de energia elétrica: análise exploratória. In: CONGRESO DE LA ASOCIACIÓN LATINOAMERICANA DE POBLACIÓN, 6., 2014, Lima, Perú. **Anais…** [*S. l.*]: ALAP, 2014.

CURRAN, S. R.; DE SHERBININ, A. Completing the picture: the challenges of bringing “consumption” into the population–environment equation. **Population and Environment,** NeW York, NY, v. 26, n. 2, p. 107-131, 2004.

DE KERF, J. L. F. The interpolation method of Sprague-Karup. **Journal of Computational and Applied Mathematics,** Amsterdam, v. 1, n. 2, p. 101-110, 1975.

DE VOS, S.; PALLONI, A. Formal models and methods for the analysis of kinship and household organization. **Population Index,** Princeton, NJ, v. 55, n. 2, p. 174-198, 1989.

DOWD, J. B. et al. Demographic science aids in understanding the spread and fatality rates of COVID-19. **Proceedings of the National Academy of Sciences,** Washington, DC, v. 117, n. 18, p. 9696-9698, 2020.

DUKHOVNOV, D.; ZAGHENI, E. Who takes care of whom in the United States? Time transfers by age and sex. **Population and Development Review,** New York, NY, v. 41, n. 2, p. 183-206, 2015.

DUNCAN, B.; TREJO, S. J. Ethnic identification, intermarriage, and unmeasured progress by Mexican Americans. In: BORJAS, G. J. (ed.). **Mexican immigration to the United States.** Chicago: University of Chicago Press, 2007. p. 229-267.

DURKHEIM, E. Introduction á la sociologie de la famille. **Annales de la Faculté des Lettres de Bordeaux,** [*S. l.*], v. 10, p. 257-281, 1888.

EDWARDS, R. D. et al. **Key equations in the Tuljapurkar-Lee model of the social security system.** Ann Arbor, MI: Michigan Retirement Research Center Research, 2003. (Working Paper, 44).

ELDER, G. H. Family history and the life course. In: HAREVEN, T. K. (ed.). **Transitions:** the family and the ife course in historical perspective. New York, NY: Academic, 1978. p. 17-64.

ELECTROBRÁS. **Projeção do nº de domicílios pela metodologia da taxa de chefia.** Rio de Janeiro, RJ, 2007.

ENGLAND. Department for Communities and Local Government. **Household interim projections:** 2011 to 2021. 2013. Available from: www.gov.uk/government/publications/household-interim-projections-2011-to-2021-in-england.

EPE – EMPRESA DE PESQUISA ENERGÉTICA. **Projeção demográfica e de domicílios.** Rio de Janeiro, RJ, 2004.

ESRI – ENVIRONMENTAL SYSTEMS RESEARCH INSTITUTE. **Evaluating population projections:** the importance of accurate forecasting. EUA, 2007. (ESRI White Paper). Available from: http://www.esri.com/library/whitepapers/pdfs/evaluating-population.pdf.

FENG, Q. et al. Forecast households at the county level: An application of the profamy extended cohort-component method in six counties of southern California, 2010 to 2040. **Population Research and Policy Review,** Amsterdam, v. 39, p. 253-281, 2020.

FIORAVANTE, E. F. **Projeção de domicílios por modelo multiestado e aplicação para previsão da frota de automóveis em Belo Horizonte.** 2009. 278f. Tese (Doutorado) – Faculdade de Ciências Econômicas, Cedeplar, Universidade Federal de Minas Gerais, Belo Horizonte, MG, 2009.

FOSDICK, B. K.; RAFTERY, A. E. Regional probabilistic fertility forecasting by modeling between-country correlations. **Demographic Research,** Germany, v. 30, n. 35, p. 1011-1034, 2014.

FREIRE, F. H. M. A. et al. Casamento e re-casamento: uma análise multivariada do mercado matrimonial no Nordeste. In: ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS, 15., 2006, Caxambu, MG. **Anais…** Belo Horizonte, MG: ABEP, 2006.

FREIRE, F. H. M. A.; AGUIRRE, M. A. C. Dinâmica entre os estados conjugais da população brasileira: uma aplicação de tábuas multi-estado para medir probabilidades de transição. In: ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS, 12., 2000, Caxambu, MG. **Anais...** Belo Horizonte, MG: ABEP, 2000.

FREITAS, R. M. V. Tendências recentes dos casamentos legais no Estado de São Paulo. **SP Demográfico – Resenha de Estatísticas Vitais do Estado de São Paulo,** São Paulo, SP, Ano 19, n. 2, 2019.

FRIAS, L. A.; RODRIGUES, P. **Brasil:** tábuas-modelo de mortalidade e populações estáveis. Rio de Janeiro, RJ: IBGE, 1981.

FRIES, J. F. The compression of morbidity. **Milbank Memorial Fund Quarterly,** New York, NY, v. 83, n. 4, p. 801-823, 1983.

FUKAWA, T. Household projection and its application to health/long-term care expenditures in Japan using INAHSIM-II. **Social Science Computer Review,** Durham, NC, v. 29, n. 1, p. 52-66, 2011.

FUNDAÇÃO SEADE. **Sistema de Projeções Populacionais para os municípios do Estado de São Paulo.** São Paulo, SP, 2017.

GEORGE, L. K.; GOLD, D. T. Life course perspectives on intergenerational and generational connections. **Marriage & Family Review,** New York, NY,v. 16, n. 1-2, p. 67-88, 1991.

GIVISIEZ, G. H. N. G.; OLIVEIRA, E. **Demanda futura por moradias:** demografia, habitação e mercado. Brasília, DF: Ministério das Cidades, 2018.

GIVISIEZ, G. H. N; OLIVEIRA, E. L. Projeção de domicílios em pequenas áreas: o uso de modelagem matemática para as taxas de chefia domiciliar. In: ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS, 19., 2014, São Pedro, SP. **Anais...** Belo Horizonte, MG: ABEP, 2014.

GIVISIEZ, G. H. N.; OLIVEIRA E. L. **Demanda demográfica até 2023 por idade e sexo para Brasil, UF e municípios segundo Tamanho do domicílio, Categorias de renda, Categorias de escolaridade do Chefe e Categorias de Déficit Habitacional.** Belo Horizonte, MG: Convênio Cedeplar, Secretaria de Habitação do Ministério das Cidades, 2011.

GIVISIEZ, G. H. N.; SAWYER, D. R. T. O; RIOS-NETO, E. L. G. Projeção da demanda demográfica por domicílios: aplicação da metodologia das taxas de chefia baseado em modelos de idade-período-coorte. In: GUIMARÃES, J. R. (org.). **Demografia dos negócios.** Campinas, SP: ABEP, 2006. p. 105-225.

GLICK, P. C. Updating the life cycle of the family. **Journal of Marriage and the Family,** US, v. 39, n. 1, p. 5-13, 1977.

GOLDANI, A. M. **Women's transitions:** the intersection of female life course, family and demographic transition in twentieth century Brazil. 1989. Tese (Doutorado) – University of Texas at Austin, Austin, 1989.

GOLDSCHEIDER, F. K.; DAVANZO, J. Semiautonomy and leaving home in early adulthood. **Social Forces,** US, v. 65, n. 1, p. 187-201, 1986.

GOLDSTEIN, J.; LEE, R. D. Demographic perspectives on the mortality of COVID-19 and other epidemics. **Proceedings of the National Academy of Sciences,** Washington, DC, v. 117, n. 36, p. 22035-22041, 2020.

GOLDSTEIN, J.; MASON, C.; ZAGHENI, E. Can grandma help with the kids? A demographic analysis of the sandwich generation. In: ANNUAL MEETING OF THE POPULATION ASSOCIATION OF AMERICA – PAA, 2011, Washington, DC. **Anais…** Alexandria, Virgínia: PAA, 2011.

GOLDSTEIN, J.; SOBOTKA, T.; JASILIONIENE, A. The end of “lowest-low” fertility? **Population and Development Review,** New York, NY, v. 35, n. 4, p. 663-699, 2009.

GONZAGA, M. R. **Compressão da mortalidade:** entendendo a variabilidade da idade à morte na população do Estado de São Paulo, 1980-2005. 2008. 124f. Dissertação (Mestrado) – Faculdade de Ciências Econômicas, Cedeplar, Universidade Federal de Minas Gerais, Belo Horizonte, MG, 2008.

GONZÁLEZ, L. M.; TORRES, E. Estimaciones de población en áreas menores en América Latina: revisión de métodos utilizados. In: CAVENAGHI, S. (org.). **Estimaciones y proyecciones de población en América Latina:** desafíos de una agenda pendiente. Rio de Janeiro, RJ: ALAP, 2012. p. 105-137. (Serie e-Investigaciones, n. 2).

GOODE, W. The sociology of the family: horizons in family theory. In: MERTON, R. K.; BROOM, L.; COTTRELL JR., L. S. (ed.). **Sociology today:** problems and prospects. New York, NY: Basic Books, 1959.

GREBENIK, E.; MACKENSEN, R. **Later phases of the family cycle:** demographic aspects. Oxford: Oxford University Press, 1989.

GUADIÁN, G. P. **Efectos esperados de la estructura y composición de los hogares en el consumo energético en México:** un pronóstico al 2030. 2016. 114f. Tese (Doutorado) – Centro de Estudios Demográficos, Urbanos y Ambientales, El Colegio de México, México, DF, 2016.

GU, D. et al. Recommendation to consider the crucial impacts of changes in smaller household size and its structure on sustainable development goals. In: UNITED NATIONS. **Crowdsourced briefs of the UN global sustainable development report.** New York, NY, 2015. Available from: https://sustainabledevelopment.un.org/topics/science/crowdsourcedbriefs.

HAMMEL, E. et al. The kin of the aged in A.D. 2000: the chickens come home to roost. In: KIESLER, S. B.; MORGAN, J. N.; OPPENHEIMER, V. K. (ed.). **Aging:** social change. New York, NY: Academic Press, 1981.

HANSEN, J. Z.; STEPHENSEN, P.; KRISTENSEN, J. B. Modeling household formation and housing demand in denmark using the dynamic microsimulation model smile. In: DREAM GENERAL CONFERENCE OF THE INTERNATIONAL MICROSIMULATION ASSOCIATION, 4., 2013, Canberra. **Anais…** Luxembourg: IMA, 2013.

HAREVEN, T. K. Aging and generational relations: a historical and life course perspective. **Annual Review of Sociology,** US, v. 20, n. 1, p. 437-461, 1994.

HILL, K.; CHOI, Y. Death distribution methods for estimating adult mortality: sensitivity with simulated data errors. In: ADULT MORTALITY IN DEVELOPING COUNTRIES WORKSHOP, 2004, Berkeley, California. **Anais…** Baltimore: Johns Hopkins University, Available from: http://www.ceda.berkeley.edu/events/AMDC_Papers/Hill_Choi_DeathDistamdc.pdf.

HOHN, C. The family life Cycle: needed extensions of the concept. In: BONGAARTS, J.; BURCH, T. K.; WACHTER, K. W. (ed.). **Family demography methods and their applications.** Oxford: Clarendon Paperbacks, 1983. p. 65-80.

HORIUCHI, S. **Healthy and mortality issues of global concern:** proceedings of the symposium on health and mortality, Brussels 19-22 nov. 1997. New York, NY: United Nations, 1999.

HYNDMAN, R. J.; ULLAH, M. S. Robust forecasting of mortality and fertility rates: a functional data approach. **Computational Statistics & Data Analysis,** Amsterdam, v. 51, n. 10, p. 4942-4956, 2007.

IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Projeções da população:** Brasil e Unidades da Federação: revisão 2018. Rio de Janeiro, RJ, 2018.

IMHOFF, E. V. et al. (ed.). **Household demography and household modeling.** New York, NY: Springer, 1995.

IRONMONGER, D. S.; AITKEN, C. K.; ERBAS, B. Economies of scale in energy use in adult-only households. **Energy Economics,** Inglaterra, v. 17, n. 4, p. 301-310, 1995.

JANSSEN, F.; KUNST, A. E. Cohort patterns in mortality trends among the elderly in seven European countries, 1950–99. **International Journal of Epidemiology,** Oxford, v. 34, n. 5, p. 1149-1159, 2005.

JANNUZZI, P. M. Proyecciones de población y políticas públicas: importancia y desafíos de las nuevas agendas. In: CAVENAGHI, S. (org.). **Estimaciones y proyecciones de población en América Latina:** desafíos de una agenda pendiente. Rio de Janeiro, RJ: ALAP, 2012. p. 87-104. (Serie e-Investigaciones, n. 2).

KANNISTO, V. et al. Reduction in mortality at advanced ages: several decades of evidence from 27 countries. **Population and Development Review,** New York, NY, v. 20, n. 4, p. 793-810, 1994.

KASHNITSKY, I.; ABURTO, J. M. **The pandemic threatens aged rural regions most.** Charlottesville, VA: Center for Open Science, 2020.

KEILMAN, N. Family projection methods: a review. In: SCHOEN, R. (ed.). **Analytical family demography.** Cham: Springer, 2018. p. 277-301.

KEILMAN, N. The threat of small households. **Nature,** London, v. 421, p. 489-490, 2003.

KHAN, H. T. A.; LUTZ, W. How well did past UN population projections anticipate demographic trends in six South-east Asian countries? **Asian Population Studies,** London, v. 4, n. 1, p. 77-95, 2008.

KNIPSCHEER, C. P. Temporal embeddedness and aging within the multigenerational family: the case of grandparenting. In: BIRREN, J. E.; BENGSTON, V. L. (ed.). **Emergent theories of aging.** [*S. l.*]: Springer Publishing Company, 1988. p. 427-446.

KONO, S. The headship rate method for projecting households. In: BONGAARTS, J.; BURCH, T. K.; WACHTER. K. W. (ed.). **Family demography**: methods and their applications. Oxford: Clarendon Press, 1987. p. 287-308.

LE PLAY, F. **L´organization de la Famille Suivant le Vrai Modéle Signalé par l´Histoire de toutes les races et de tous les temps.** Paris: Tequi, 1871.

LEE, R. D. Probabilistic approaches to population forecasting. In: LUTZ, W.; VAUPEL, J. W.; AHLBURG, D. A. **Frontiers of population forecasting.** New York, NY: Population Council, 1999.

LEE, R. D.; CARTER, L. R. Modeling and forecasting US mortality. **Journal of the American Statistical Association,** New York, NY, v. 87, n. 419, p. 659-671, 1992.

LEIWEN, J.; O'NEILL, B. Toward a new model for probabilistic household forecasts. **International Statistical Review,** Oxford, v. 72, n. 1, p. 51-64, 2004.

LENZEN, M.; MURRAY, S. A. A modified ecological footprint method and its application to Australia. **Ecological Economics,** Amsterdam, v. 37, n. 2, p. 229-255, 2001.

LEON, N.; MOREIRA, J. Projeções da população e do nº de domicílios: uma incerteza para a projeção do mercado de energia elétrica em um horizonte de dez anos? In: WORKSHOP DEMOGRAFIA DOS NEGÓCIOS, 2005, Salvador, BA. **Anais...** Belo Horizonte, MG: ABEP, 2005.

LESTHAEGHE, R.; VAN DE KAA, D. J. **Two demographic transitions. Population growth and decline.** 1986. p. 9-24.

LI, N.; LEE, R.; GERLAND, P. Extending the Lee-Carter method to model the rotation of age patterns of mortality decline for long-term projections. **Demography,** New York, NY, v. 50, n. 6, p. 2037-2051, 2013.

LIDDLE, B. Consumption-driven environmental impact and age structure change in OECD countries: a cointegration-STIRPAT analysis. **Demographic Research,** Germany, v. 24, n. 30, p. 749-770, 2011.

LIMA, E. E. C. et al. **Exploring excess mortality in Latin America in the context of covid pandemic:** the cases of Brazil and Ecuador. Charlottesville, VA: Center for Open Science, 2020.

LIMA, E. E. C.; TOMÁS, M. C.; QUEIROZ, B. L. The sandwich generation in Brazil: demographic determinants and implications. **Revista Latinoamericana de Población,** México, DF, v. 9, n. 16, p. 59-73, 2015.

LIU, J. et al. Effects of household dynamics on resource consumption and biodiversity. **Nature,** London, v. 421, p. 530-533, 2003.

LUTZ, W.; SKIRBEKK, V.; TESTA, M. R. The low-fertility trap hypothesis: forces that may lead to further postponement and fewer births in Europe. **Vienna Yearbook of Population Research,** Vienna, Austria, p. 167-192, 2006.

LUY, M; MINIGUAWA, Y. Gender gaps: life expectancy and proportion of life in poor health. **Health Reports,** [*S. l.*], v. 25, n. 12, p. 12-19, 2014.

MARCOS, M.; MÓDENES, J. A. **Proyecciones de hogares y demanda de vivienda en la Región Metropolitana de Buenos Aires.** Bellaterra: Centre d'Estudis Demogràfics, 2019. (Papers de Demografia, 464).

MARTINS, R. N. **A adequação da habitação de interesse social aos arranjos familiares no município de São Paulo em 2010.** 2019. 190f. Dissertação (Mestrado) – Instituto de Filosofia e Ciências Sociais, Universidade Estadual de Campinas, Campinas, SP, 2019.

MASON, A.; RACELIS, R. A comparison of four methods for projecting households. **International Journal of Forecasting,** Amsterdam, v. 8, n. 3, p. 509-527, 1992.

MASON, C.; ZAGHENI, E. The sandwich generation: demographic determinants of global trends. In*:* ANNUAL MEETING OF THE POPULATION ASSOCIATION OF AMERICA – PAA, 2014, Boston. **Anais…** Alexandria, Virgínia: PAA, 2014.

MATEA, M. L. **La demanda potencial de vivienda principal.** Madrid: Banco de España, 2015. (Documentos Ocasionales, n. 1504).

MCDONALD, P. Gender equity in theories of fertility transition. **Population and Development Review,** New York, NY, v. 26, n. 3, p. 427-439, 2000.

MENG, X.; GREGORY, R. G. Intermarriage and the economic assimilation of immigrants. **Journal of Labor Economics,** Chicago, v. 23, n. 1, p. 135-174, 2005.

MODELL, J.; FURSTENBERG JR., F. F.; STRONG, D. The timing of marriage in the transition to adulthood: continuity and change, 1860-1975. **American Journal of Sociology,** Chicago, v. 84, (Supl.), p. S120-S150, 1978.

MOURA, E. C. et al. Desigualdades de gênero na mortalidade por causas externas no Brasil, 2010. **Ciência & Saúde Coletiva,** Rio de Janeiro, RJ, v. 20, n. 3, p. 779-788, 2015.

MURPHY, M. Family and kinship networks in the context of ageing societies. In: TULJAPURKAR, S.; OGAWA, N.; GAUTHIER, A. H. (ed.). **Ageing in advanced industrial states:** riding the age waves. Berlin: Springer, v. 3, 2010. p. 263-285. (International Studies in Population Book, 8).

MURPHY, M. Modelling households: a synthesis. In: MURPHY, M. J.; HOBCRAFT, J. (ed.). **Population research in Britain.** London: Population Investigation Committee; London School of Economics, 1991. p. 157-176 (A supplement to Population Studies, v. 45).

MURRAY, C. J. L. et al. Modified logit life table system: principles, empirical validation, and application. **Population Studies,** London, v. 57, n. 2, p. 165-182, 2003.

MYERS, R. J. Errors and bias in the reporting of ages in census data. **Transactions of the Actuarial Society of America,** Schaumburg, v. 41, part 2, n. 104, p. 395-415, 1940.

MYRSKYLÄ, M.; KOHLER, H.; BILLARI, F. **High development and fertility:** fertility at older reproductive ages and gender equality explain the positive link. Germany: Max-Planck-Institut für demografi sche Forschung; Max Planck Institute for Demographic Research, 2011.

NEPOMUCENO, M. R. et al. Besides population age structure, health and other demographic factors can contribute to understanding the COVID-19 burden. **Proceedings of the National Academy of Sciences,** Washington, DC, v. 117, n. 25, p. 13881-13883, 2020.

NERI, A. L. (org.). **Cuidar de idosos no contexto da família:** questões psicológicas e sociais. 3. ed. Campinas, SP: Alínea, 2012.

NERI, A. L.; SOMMERHALDER, C. As várias faces do cuidado e do bem-estar do cuidador. In: NERI, A. L. (org.). **Cuidar de idosos no contexto da família:** questões psicológicas e sociais. Campinas, SP: Alínea, 2002. p. 9-64.

OEPPEN, J.; VAUPEL, J. W. Broken Limits to Life Expectancy. **Science,** Washington, DC, v. 296, n. 5570, p. 1029-1031, 2002.

OLIVEIRA, J.; FERNANDES, F. Metodologia e considerações acerca da projeção da população do Brasil: 1980-2020. **São Paulo em Perspectiva,** São Paulo, SP, v. 10, n. 2, p. 116-123, 1996.

OLIVEIRA, S.; SABOIA, A. L.; SOARES, B. **Dimensões preliminares da responsabilidade feminina pelos domicílios:** um estudo do fenômeno a partir dos censos demográficos 1991 e 2000. Rio de Janeiro, RJ: IBGE, 2002. (Série Textos para Discussão, n. 7).

OLSHANSKY, S. J. et al. A potential decline in life expectancy in the United States in the 21st century. **New England Journal of Medicine,** Waltham, MA, v. 352, n. 11, p. 1138-1145, 2005.

O'NEILL, B. C.; CHEN, B. S. Demographic determinants of household energy use in the United States. **Population and Development Review,** New York, NY, v. 28, (Supl.), p. S53-S88, 2002.

PAES, N. A. Avaliação da cobertura dos registros de óbitos dos Estados brasileiros em 2000. **Revista de Saúde Pública,** São Paulo, SP, v. 39, n. 6, p. 882-890, 2005.

PARSONS, T. The social structure of the family. In: ANSHEN, R. N. **The family: its function and destiny.** New York, NY: Harper and Brothers, 1949. p. 173-201.

PASCARIU, M. D. et al. **Introduction to ungroup.** [*S. l.: s. n.*], 2018.

PRIOUX, F. (ed). **La famille dans le pays développés:** permanences et changements: actes du séminaire sur les nouvelles formes de vie familiale dans les pays développés (Congrès et colloques). Paris: INED, 1990.

PRSKAWETZ, A.; LEIWEN, J.; O'NEILL, B. C. Demographic composition and projections of car use in Austria. **Vienna Yearbook of Population Research,** Vienna, Austria, v. 2, n. 1, p. 175-202, 2004.

QUILODRÁN, J. ¿Un modelo de nupcialidad postransicional en América Latina? In: CONGRESSO DA ASSOCIAÇÃO LATINOAMERICANA DE POPULAÇÃO – ALAP, 3., 2008, Córdoba, Argentina. **Anais...** [*S. l.*]: ALAP, 2008.

RAMSAY, J. O.; SILVERMAN, B. W. Principal components analysis for functional data. In: RAMSAY, J. O.; SILVERMAN, B. W. **Functional data analysis.** New York, NY: Springer, 2005. p. 147-172.

RIZZI, S.; GAMPE, J.; EILERS, P. H. C. Efficient estimation of smooth distributions from coarsely grouped data. **American Journal of Epidemiology,** US, v. 182, n. 2, p. 138-147, 2015.

ROGERS, A. Parameterized multistate population dynamics and projections. **Journal of the American Statistical Association,** New York, NY, v. 81, n. 393, p. 48-61, 1986.

RUGGLES, S. The decline of intergenerational coresidence in the United States, 1850 to 2000. **American Sociological Review,** US, v. 72, n. 6, p. 964-989, 2007.

RUGGLES, S. Confessions of a microsimulator: problems in modeling the demography of kinship. **Historical Methods: A Journal of Quantitative and Interdisciplinary History,** London, v. 26, n. 4, p. 161-169, 1993.

RYDER, N. B. Recent developments in the formal demography of the family. In: INTERNATIONAL POPULATION CONFERENCE, 1985, Florence. **Anais...** França: IUSSP, 1985.

SABÓIA, A. L.; SOARES, C. O conceito de chefia nas pesquisas domiciliares através do recorte por sexo e presença do cônjuge-uma contribuição à discussão da “feminização da pobreza”. **Revista Gênero,** Niterói, RJ, v. 4, n. 2, p. 53-71, 2012.

SÁNCHEZ PEÑA, L. Hogares y consumo energético en México. **Revista Digital Universitaria,** México, DF, v. 13, n. 10, p. 1-8, 2012.

SETTERSTEN JR, R. A.; FURSTENBERG, F. F.; RUMBAUT, R. G. (ed.). **On the frontier of adulthood:** theory, research, and public policy. Chicago: University of Chicago Press, 2008.

SHRYOCK, H. S.; SIEGEL, J. S. **The methods and materials of demography.** San Diego: Academic Press, 1976.

SILVA, C. C. **População e consumo:** efeitos de características sociodemográficas sobre o consumo de energia elétrica domiciliar em Lucas do Rio Verde (MT) e Santarém (PA). 2013. 261f. Thesis (PhD in Demography) – Faculty of Philosophy, History and Social Sciences, University of Campinas, Campinas, SP, 2013.

SLOGGETT, A. Measuring fertility. In: IUSSP. **Population analysis for policy and programmes.** França, 2015. Available from : http://papp.iussp.org/sessions/papp101_s04/PAPP101_s04_010_010.html. Acesso em: 02 abr. 2019.

SMITH, S. K.; TAYMAN, J.; SWANSON, D. A. **State and local population projections:** methodology and analysis. Germany: Springer Science & Business Media, 2006.

SOKOLOWSKY, J. **Living arrangements of older persons and family support in less developed countries.** Florida: University of South Florida, Bayboro Campus, 2008.

STATISTICS KOREA. **Household projections for Korea by Province (2017 ~ 2047).** Republic of Korea, s.d. Available from: http://kostat.go.kr/portal/eng/pressReleases/8/2/index.board.

THERBORN, G. **Between sex and power:** family in the world 1900-2000. London: Routledge, 2004.

TOCQUEVILLE, A. **Democracy in America.** Chicago: University of Chicago Press, 1969.

TOMÁS, M. C.; LIMA, E. E. C.; QUEIROZ, B. L. The myth of the sandwich generation in Brazil le mythe de la génération sandwich au Brésil. **NIUSSP,** july 2018. Available from : https://www.niussp.org/article/the-myth-of-the-sandwich-generation-in-brazil/.

UNFPA – UNITED NATIONS POPULATION FUND. **Population projections:** concepts and methods. New York, NY, Available from : http://papp.iussp.org/sessions/papp101_s10/PAPP101_s10_010_010.html.

UNITED NATIONS. Department of Economic and Social Affairs. Population Division. **Household size and composition 2018.** New York, NY, 2018.

UNITED NATIONS. Departamento de Asuntos Económicos y Sociales. División de Estadística. **Principios y recomendaciones para los censos de población y vivienda:** revision 2. New York, NY, 2007.

UNITED NATIONS. Department of Economic and Social Affairs. Statistics Division. **Principles and recommendations for population and housing censuses.** New York, NY, 1998

UNITED NATIONS. Department of International Economic. **Unabridged model life Tables corresponding to the new United Nations model life Tables for developing countries.** New York, NY, 1982.

UNITED NATIONS. Department of Economic and Social Affairs. Methods of projecting households and families. In: UNITED NATIONS. **Manual VII.** New York, NY, 1973. (Population Studies, n. 54).

VAN IMHOFF, E.; KEILMAN, N. **LIPRO 2.0:** an application of a dynamic demographic projection model to household structure in the Netherlands. Amsterdam: Swets & Zeitlinger, 1991.

VIEIRA, J. M. Os domicílios de responsabilidade compartilhada no Brasil em 2010: uma questão de homogamia marital? In: Congreso Latinoamericano de Población – ALAP, 7.,

ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS, 20., 2017, Foz do Iguaçu, PR. **Anais...** Belo Horizonte, MG: ABEP, 2017.

VIEIRA, J. M.; ALVES, L. C. O comportamento da idade média à união e ao casamento no Brasil em 2000 e 2010. **Revista Latinoamericana de Población,** México, DF, v. 10, n. 19, p. 107-125, 2016.

VIEIRA, J. M. **Transição para a vida adulta em São Paulo:** cenários e tendências sóciodemográficas. 2009. 208f. Tese (Doutorado) – Instituto de Filosofia e Ciências Humanas, Universidade Estadual de Campinas, Campinas, SP, 2009.

WAG – WELSH ASSEMBLY GOVERNMENT. **Household projections across the United Kingdom:** technical report. United Kingdom, 2011.

WAJNMAN, S. **Demografia das famílias e dos domicílios brasileiros.** Belo Horizonte, MG: UFMG/FACE, 2012.

WALDVOGEL, B. C.; CAPASSI, R.; MORAIS, L. C. C. Qual trajetória a população paulista seguirá no futuro? **SP Demográfico – Resenha de Estatísticas Vitais do Estado de São Paulo,** São Paulo, SP, Ano 18, n. 4, 2018.

WESTERMARCK, E. **Human marriage in the history of mankind.** São Paulo, SP: Macmillan, 1891.

WHELPTON, P. K. An empirical method for calculating future population. **Journal of the American Statistical Association,** New York, NY, v. 31, n. 195, p. 457-473, 1936.

WILLEKENS, F.; PUTTER, H. Software for multistate analysis. **Demographic Research,** Germany, v. 31, n. 14, p. 381-420, 2014.

WILLEKENS, F. Family and household demography. In: UNESCO. **Encyclopedia of Life Support Systems.** Paris, v. 2, 2010. p. 86-112.

WILMOTH, J. R.; HORIUCHI, S. Do the oldest old grow old more slowly? In: ROBINE, J. M. et al. (ed.). **The paradoxes of longevity.** Berlin: Springer, 1999. p. 35-60.

WOLF, D. Coresidence with an aged parent: lifetime patterns and sensitivity to demographic changes. In: UNITED NATIONS CONFERENCE ON POPULATION AGING IN THE CONTEXT OF FAMILY, 1990, Kitakyushu, Japan. **Anais…** New York, NY: United Nations, 1990.

XENOS, P.; KONO, S. **Thinking about family change.** Ordina: Liège, 1992.

YÁNEZ PETER, N. L. **Proyecciones de hogares en el Ecuador por su tamaño, mediante métodos de proyección de proporciones condicionales respecto al número de miembros de hogar, a nivel provincial y nacional.** Quito, 2015. Available from : http://bibdigital.epn.edu.ec/handle/15000/9260.

YEPEZ MARTÍNEZ, B. Y. **Proyecciones de hogares:** una aplicación para Venezuela al horizonte 2021. 2010. Tese (Doutorado) – Univesitat Autónoma de Barcelona – UAB, 2010.

YOUSIF, H. M. Population, biomass and the environment in Central Sudan. **International Journal of Sustainable Development & World Ecology,** Inglaterra, v. 2, n. 1, p. 54-69, 1995.

ZAGHENI, E. The leverage of demographic dynamics on carbon dioxide emissions: does age structure matter? **Demography,** New York, NY, v. 48, n. 1, p. 371-399, 2011.

ZENG, Y. et al. Effects of changes in household structure on future housing demand in Hebei Province, China. **Genus,** Roma, v. 69, n. 2, p. 85-111, 2014a.

ZENG, Y. et al. **Household and living arrangement projections:** the extended cohort-component method and applications to the U.S. and China. New York, NY: Springer Netherlands, 2014b.

ZENG, Y. **Household and population projections at sub-national levels:** an extended cohort-component approach. Germany: Max-Planck-Institut für demografi sche Forschung; Max Planck Institute for Demographic Research, 2010.

ZENG Y. et al. U.S. Family household momentum and dynamics: extension of profamy method and application. **Population Research and Policy Review,** Amsterdam, v. 25, n. 1, p. 1-41, 2006.

ZENG, Y.; VAUPEL, J. W.; WANG, Z. Household projection using conventional demographic data. **Population and Development Review,** New York, NY, v. 24, (Supl.), p. 59-87, 1998.

ZENG Y., VAUPEL J. W., WANG Z. A multi-dimensional model for projecting family households: with an illustrative numerical application. **Mathematical Population Studies,** London, v. 6, n. 3, p. 187-216, 1997.

ZENG, Y. **Family dynamics in China:** a life table analysis. Wisconsin: The University of Wisconsin Press, 1991.

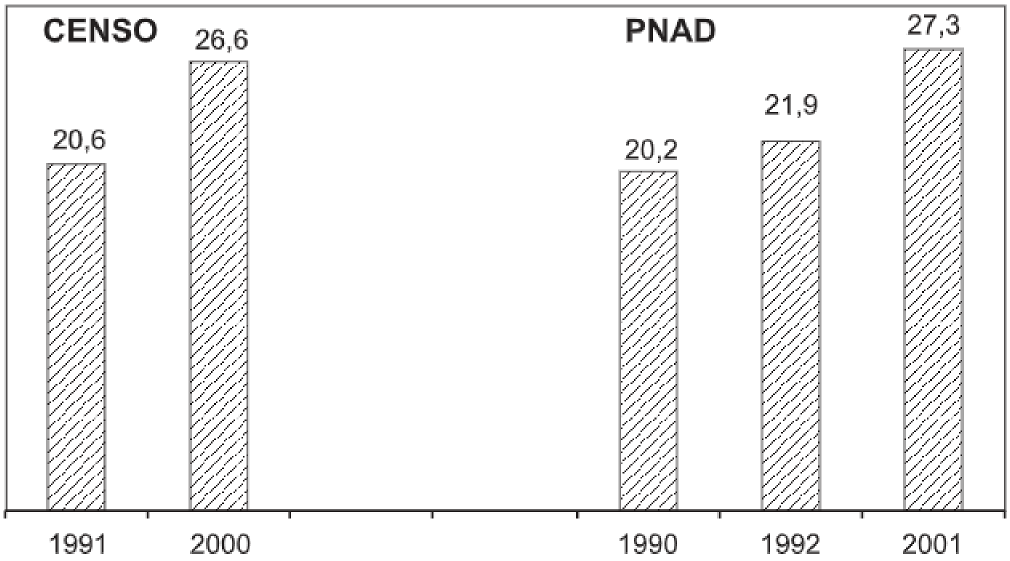
ZENG, Y. **User's manual and software of "FAMY":** a PC computer program for family status life table analysis. Groningen: iec ProGAMMA (Inter-university Expert Center for Computer Software Development and Distribution), 1990.

ZENG, Y. Changing demographic characteristics and the family status of Chinese women. **Population Studies,** London, v. 42, n. 2, p. 183-203, 1988.

ZENG, Y. Changes in family structure in China: a simulation study. **Population and Development Review,** New York, NY, v. 12, n. 4, p. 675-703, 1986.

APPENDIX

FIGURE 44 – Proportion of households with "female head" (1991, Census) and "responsible for the household" (2000 Census and PNAD), Brazil



Source: Sabóia and Soares (2012).

FIGURE 45 – Total Fertility Rate annual trend – Brazil (1920-2050)

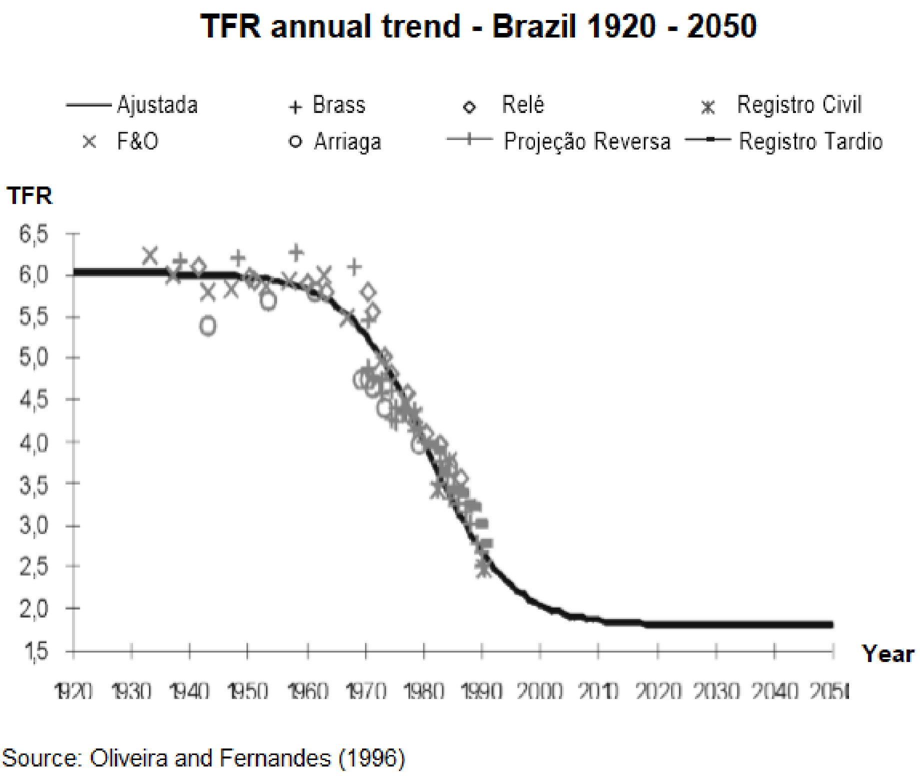
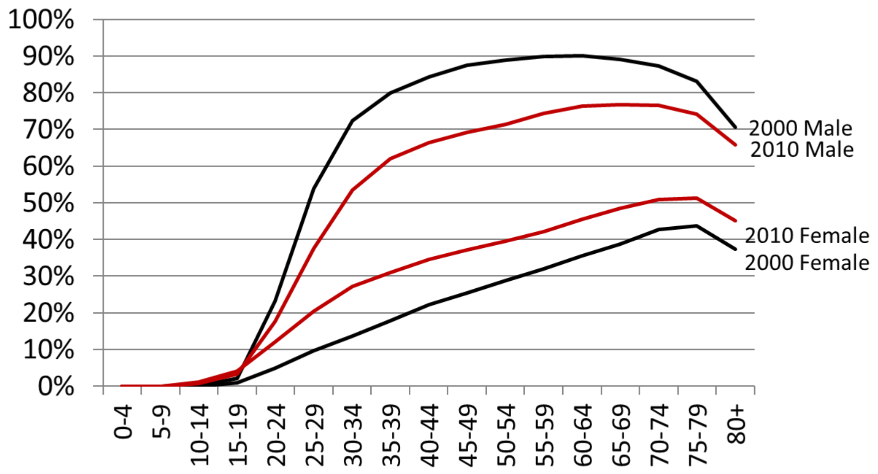


FIGURE 46 – São Paulo’s headship rates (x 100) by sex, 2000 and 2010



Source: Author’s elaboration using Census data (2000; 2010).