Subnational Fertility Projections in Brazil – a Bayesian Probabilistic Approach Application

Gabriel Mendes Borges

Abstract

The main objective of this paper is to project fertility rates for the sub-national level in Brazil using the UN's Bayesian probabilistic approach in order to compare with the most recent fertility information and with the deterministic projection released by IBGE (Brazilian Institute of Geography and Statistics). Results show that out-of sample projections systematically overestimate the TFR in 2010, since there was an acceleration of the fertility decline in the 2000s. Social advances observed in this decade might have influenced such rapid fertility decline. For projections with starting point in 2010, IBGE's estimations still present lower figures than those projected by the Bayesian model, in addition to a different convergence pattern in 2030. Bayesian approach projections offer a promising alternative to sub-national level projections in Brazil, providing point estimations and uncertainty of measures. However, some adjustments may be added to the model such as the incorporation of age-specific fertility rates and some covariates related to fertility, like educational attainment.

Introduction

Population projections are fundamental part of understanding the future of human population, playing a central role in the debate about population, development and environment. They are broadly used in planning and monitoring public policies and private decisions. In the Brazilian context, population projections and estimates are also used in funding allocations to states and municipalities, sample surveys expansions, besides serving as the denominators for the calculation of several indicators. For all of these reasons, besides providing national outcomes, demand is increasing for subnational level projections.

Fertility is the main responsible variable in modifying the population size and structure, especially in countries with low migration rates, like Brazil. Therefore, an important part of the population projection accuracy is related to fertility estimations, which constitute a challenge, since it involves high level of uncertainty.

Population projections in the Brazil, with rare exceptions, have underestimated the fertility decline since the 1970's. Official population projections for this period projected Total Fertility Rate (TFR) of 3.6 and 4.8, in lower and upper scenarios respectively, for the end of the 20th century (CBED, 1974). In the 1980s, projections estimated the TFR around 2.5 for 2010 (FIBGE and CELADE, 1984). In the 1990s, after the preliminary results of the 1991 Census, Carvalho (2004) says that even among demoGraphers fertility decline in the 1980s was higher than expected. A rare example that have underestimated fertility was the preliminary projection released by IBGE, the Brazilian Institute of Geography and Statistics (MENDES, 1994; OLIVEIRA and FERNANDES, 1996). Most recently, in the the beginning of the 2000s, most official projections performed by IBGE (IBGE, 2001; OLIVEIRA et al., 2004), CELADE (CEPAL and CELADE, 2004) and the United Nations (Desa(2003, 2006, 2007)) still estimated Brazilian fertility above the replacement level around 2010. The out-of-sample model to Brazil using the United

Nations (UN) Bayesian Probabilistic Approach also tended to under-predict the fertility decline (Raftery et al., 2009).

These examples show the unexpected velocity of fertility decline in Brazil. Despite some evidences of slightly fertility decline in the end of the 19th century and the early decades of the 20th century (Rios-Neto, 2000), fertility starts falling rapidly in the second half of the 1960s. Total Fertility Rate has reduced to more than 6 children per woman in 1960 to less than 2 in 2010, representing a reduction of more than 4 children per women in only 50 years. Although the decline has been generalized, the timing and intensity vary for each region.

Brazilian fertility transition happened along with changes in the age structure in which women have their children. Until 2000, there was an increasing concentration of births at younger ages of the reproductive period, between 15 and 24 years old, with an intensification of this trend during the 1990s. Since 2000, there has been a reversal in the tendency, with births concentration within a smaller interval at older ages, although high fertility among adolescents and young women persists.

Understanding such historic fertility trends is fundamental part of producing more accurate fertility projections, although there will always be some level of uncertainty. Despite the existence of these uncertainties and demand for some measure of variability, IBGE has released its official projections considering only one deterministic fertility scenario (IBGE, 2013). One of the reasons of doing that is the high number of legal proceedings challenging the official population estimates (Borges et al., 2011), which could increase with the incorporation of uncertainty in the estimates.

In recent years, there has been an increasing emphasis by national statistical offices to include uncertainty in their official population projections (Abel et al., 2010) although national statistical offices still use deterministic variant projections to provide uncertainty (Lutz and Goldstein, 2004)

Probabilistic approaches have been criticized in recent probabilistic work (Wilson and Rees, 2005), even regarding to the difficulty of interpretation of the results (Lee, 1998), but at the same time they appear as promising and innovative methods to project fertility, especially considering a Bayesian approach (Abel et al., 2010).

In this context, the main objective of this paper is to project the fertility rates for the Brazilian subnational levels (Major Regions and Units of Federation)[1] using a method that include uncertainty in order to compare the results with the point estimates of deterministic projections released by IBGE (2013).

There are several methods for the projecting fertility[2] and, among them, we choose a probabilistic method, which is the Bayesian approach that has been used by United Nations in the last revisions of the World Population Prospects (DeSA, 2011, 2013a).

The first specific objective of this work is to run out-of-sample Bayesian probabilistic projection considering the years 1940 to 2000 and then compare the results to the fertility rates estimated for 2010. In the light of these comparisons, we discuss the recent trend of the Brazilian fertility.

[1] Brazil is formed by five Major Regions – North (NO), Northeast (ND), Southeast (SD), South (SU) and Central-West (CO) – and 27 Units of the Federation (UF) – 26 States and the Federal District.

[2] For a more complete reference and history of methods for fertility projection see, for example, (De Beer, 1992; Booth, 2006; Wilson and Rees, 2005).

The second specific objective is to project the population using the Bayesian probabilistic projections approach from 1940 to 2030 and compare with the estimations published by IBGE (2013). In doing this, it is possible to evaluate those projections, both in terms of level and regional convergence, discussing the fertility trends for subnational levels in Brazil for the next two decades. We add some socioeconomic status measures that are normally related to fertility to evaluate the results.

We conclude by discussing how probabilistic projections, especially this Bayesian approach, can contribute to the subnational projections in Brazil, considering that it could offer a more realistic projection, since it is associated with uncertainty measures.

Data and Methods

Data

Besides difficulties in establishing methods and future hypotheses, adequacy of data is also an important issue in demoGraphic estimations in Brazil. For subnational population projections, this issue is even more problematic, since these projections depend greatly on the quality of inputs required for the cohort-component method.

Due to the difficulties in estimating and projecting demoGraphic parameters, IBGE, the institution responsible for releasing the national and subnational projections and estimates, only released its first official projections for the UFs using the cohort-component method in 2013[3] (IBGE, 2013).

The quality of the administrative records in Brazil varies across geographic regions, reflecting the diversity regarding population size and composition and the high socioeconomic inequality in the country. The most developed Units of Federation (UF) have good quality of births administrative records, showing coverage close to 100%. On the other hand, many States, especially those in the North and Northeast regions, present substantial births under-registration. Maranhão, for example, had 37% of the births not reported in 2000. Although there have been some improvements in the last decade, almost 20% of the births were still not registered in 2010 (Borges and Silva, 2015). Thus, considering the quality of the administrative records, fertility in Brazil is usually estimated by indirect techniques, using censuses and household sample surveys. Considering the lack of estimates precision of household surveys for the geographic level required, in the same way that (IBGE, 2013), only the information from the Censuses are used. The Brazilian Censuses have the strength of having information that permit measurement of fertility in every decennial Census since 1940.

The information for all countries used to perform the Bayesian probabilistic projection derives from the 2012 revision of the World Population Prospects (DeSA, 2013a).

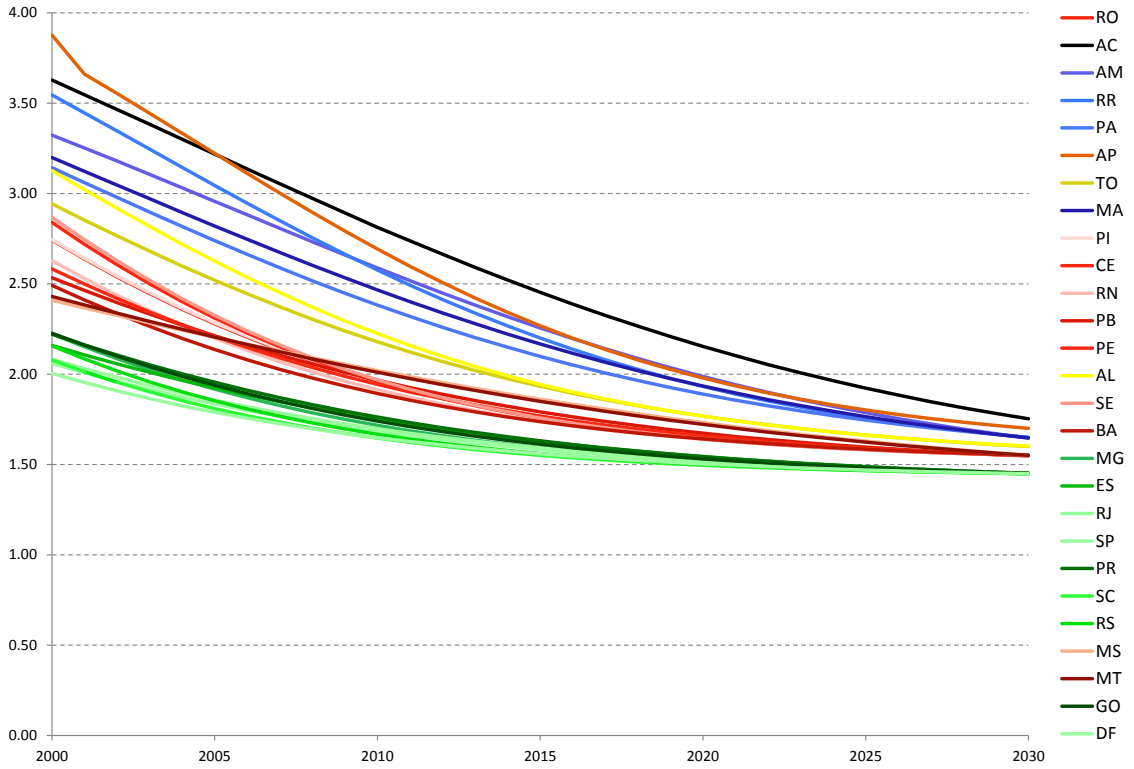
IBGE's 2013 Projection

The method used in IBGE's Projection, released in 2013, was a logistic interpolation of the Total Fertility Rates (TFR) from 2000 to 2010 and an extrapolation of these rates until 2030, considering limits that vary according to the characteristics of each UF.

[3] Prior to 2013, IBGE used the "Apportionment Method" to estimate only the total UFs' population (IBGE, 2008).

The intrinsic assumption is that fertility will keep falling, and there will be a regional, but not national, convergence in the future. Figure 1 shows the TFT estimated for the Units of Federation from 2000 to 2030. It was assumed that there will be a reduction in the difference between the UFs, fallowing the recent trend, but there will not be a complete convergence in the next two decades, due to the expected persistence of cultural differences and socioeconomic inequalities (Campos and Borges, 2015; IBGE, 2013).

Figure 1 – Total Fertility Rates by Units of Federation – 2000-2030



Source: (IBGE, 2013)

For the age-specific fertility rates, considering the relation between this variable and the women's educational level, as well as the necessity of using a national experience for these estimations, the limit fertility pattern was defined taking into account the observed pattern among the most educated women in a selected number of low fertility UFs in 2010 (Campos and Borges, 2015; IBGE, 2013).

Bayesian Probabilistic Projection

Until recently, UN used deterministic approaches to estimate future fertility, including three scenarios to incorporate uncertainty: medium, high (adding half a child to the TFR) and low (reducing half a child to the TFR).

Alkema et al. (2011) propose a Bayesian method to produce TFR projections for all countries based on the historic trend of the TRF. It includes three phases: the pre-transitional high fertility; the fertility transition from high to replacement level or below; and the post-transition low fertility, which includes recovery from below-replacement level and oscillations around replacement level.

The second phase, which is the fertility transition in which the TFR decreases from high fertility levels towards or below replacement level fertility, is the most important for the

context of this paper, since is the phase that all Brazilian regions have been experiencing in the range of time analyzed in this paper. In this phase, the method models the fertility decline as the sum of two logistic functions that depend on the current TFR level in addition to a random term[4].

Ševcıková and Gerland (2013) show the methods and the tools used by the United Nation Population Division in the preparation of their last population projections revisions through the R package bayesTFR, encouraging the application of this method also to subnational datasets. Gerland (2011) also presented one example of application of bayesTFR package to states in Brazil, showing the steps involved in preparing and running the estimations, in addition to some results.

Out-of-Sample Prediction and Recent Fertility Decline in Brazil

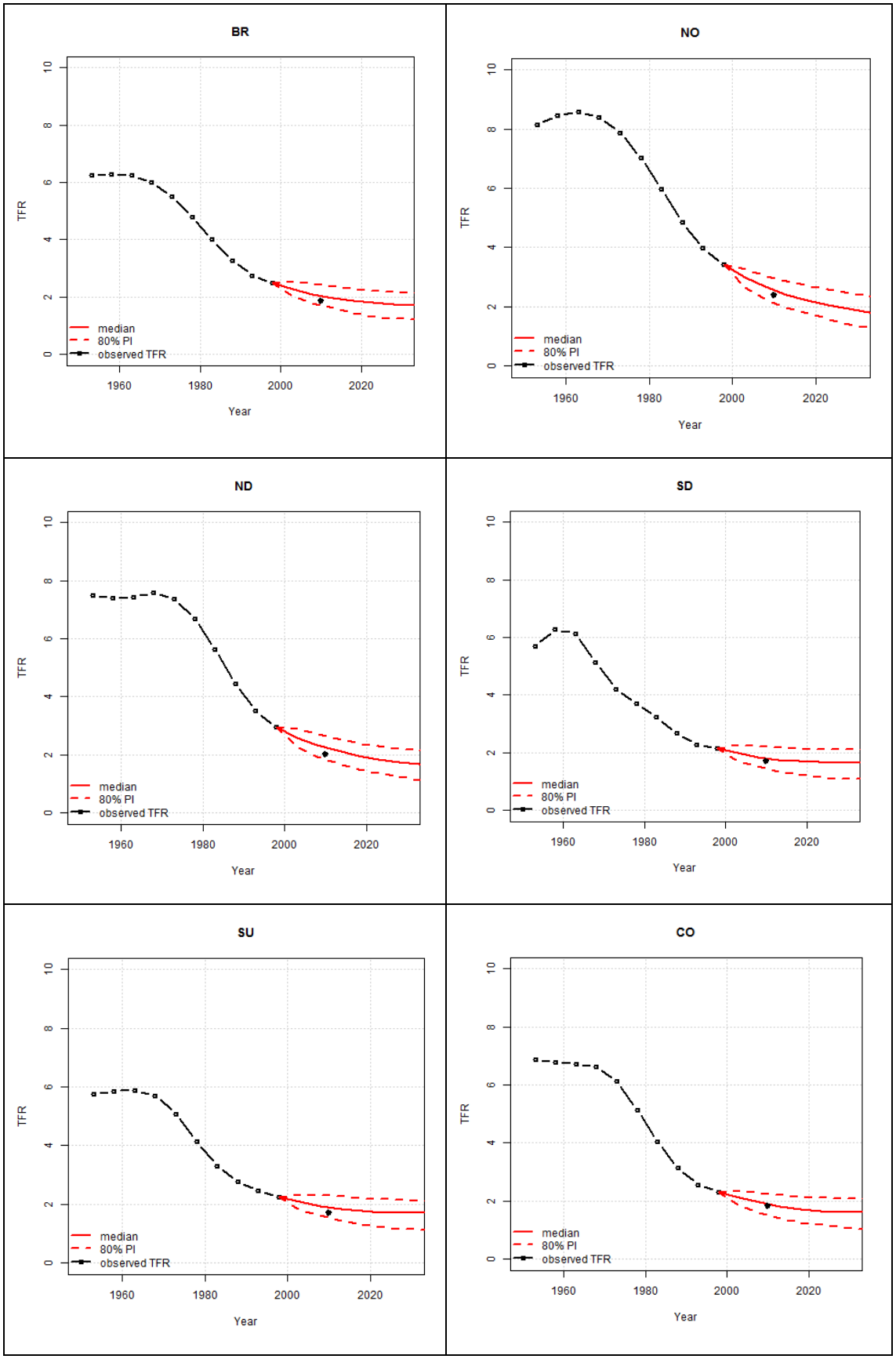
Out-of-sample predictions using fertility information for Brazil, the 5 Major Regions and the 27 UF's from 1950 and 2000 and the TFR estimated for all countries (DeSA, 2013a) are compared with the Total Fertility Rates estimated by 2010. Figure 2 shows the estimated (1950-2000) and projected (2000-2030) TFR for Brazil and its Major Regions, with a confidence interval of 80%, in addition to the TFR observed in 2010.

Projections systematically overestimate the TFR in 2010, even though the values observed in 2010 are within the confidence interval. Differences were greater in the Northeast (ND) and South (SU) Regions, where projection overestimated fertility in more than 10%. Prediction for the Central-West (CO) Region is the closest to the TFR estimated by 2010, and the difference between predicted and estimated values is 5% and 6% respectively for the Regions North and Southeast.

Confidence intervals are relatively high, even for short-term projection. The difference between the upper and lower level, even considering the 80% projection interval is more than 0.8 children for the period 2010-2015. For 2030-2035 it is greater than 1.0 for all the Major Regions. For the country as a whole, it tends to be smaller.

[4] For more details of the model, see Alkema et al. (2011) and Ševcıková (2011).

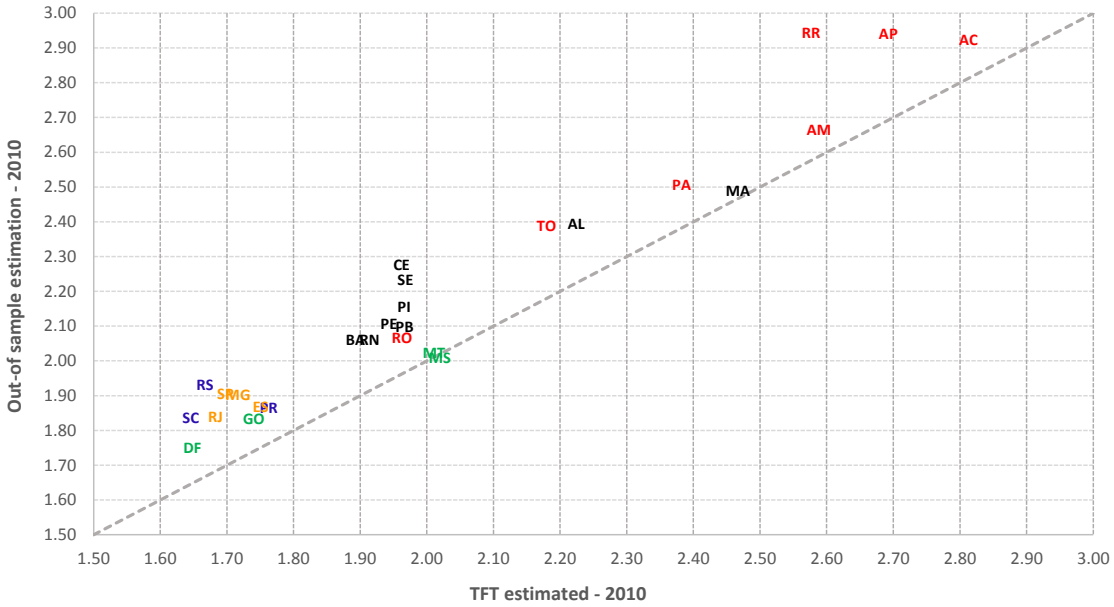
Figure 2 – Total Fertility Rates estimated (1950-2000) and projected (2000-2030) - Brazil and Major Regions



Source: (DeSA, 2013a; IBGE, 2013)

Figure 3 shows the Total Fertility Rates projected and observed for each of the Units of Federation in 2010, where it is possible to see that for the lower geographic levels there is also a tendency of the projections to overestimate the TFR in 2010. Only Maranhão (MA), Mato-Grosso (MT) e Mato-Grosso do Sul (MS) present similar figures. The greater discrepancies are in Ceará (CE), Rio Grande do Sul (RS), Roraima (RR) e Sergipe (SE).

Figure 3 – Total Fertility Rates projected and observed by Units of Federation in 2010



Source: (DeSA, 2013a; IBGE, 2013)

Differences between projected and observed fertility are not associated with either the fertility or the socioeconomic levels of the regions. It depends on the decrements in TFR according to the fertility level. The function that model the fertility implies decreasing decrements associated with lower levels of fertility, but most of the UFs actually accelerated the fertility decline in the last decade, reversing the historical trend. Even states who were already below the replacement-level presented a faster fertility decline.

In this sense, it is important to contextualize the Brazilian conjuncture during this period. From 2000 and 2010, Brazil experienced important social advances, with sharp decline in inequality and poverty in addition to improvements in educational levels, which increased more quickly during this period than had increased historically (Barros et al., 2010).

Fertility differentials by income and education are still high, but have reduced in the last decade, since fertility has declined more among the less educated and poorer women (Berquó and Cavenaghi, 2014; IBGE, 2012). Cavenaghi and Berquó (2014) attribute some of the variations between groups to diversity in the behavior and access to contraceptive methods in the population, besides associate the changes in the income and education structure to the decline in fertility.

Population Projections: 2010-2030

This section presents the results of the UN's Bayesian probabilistic approach application from 2010 to 2030 for Brazil, its Major Regions and Units of Federation.

Figure 4 shows the Bayesian estimates of the fertility decline curves for Brazil and Major Regions. They have similar shape, but each region has its particularity. One of the

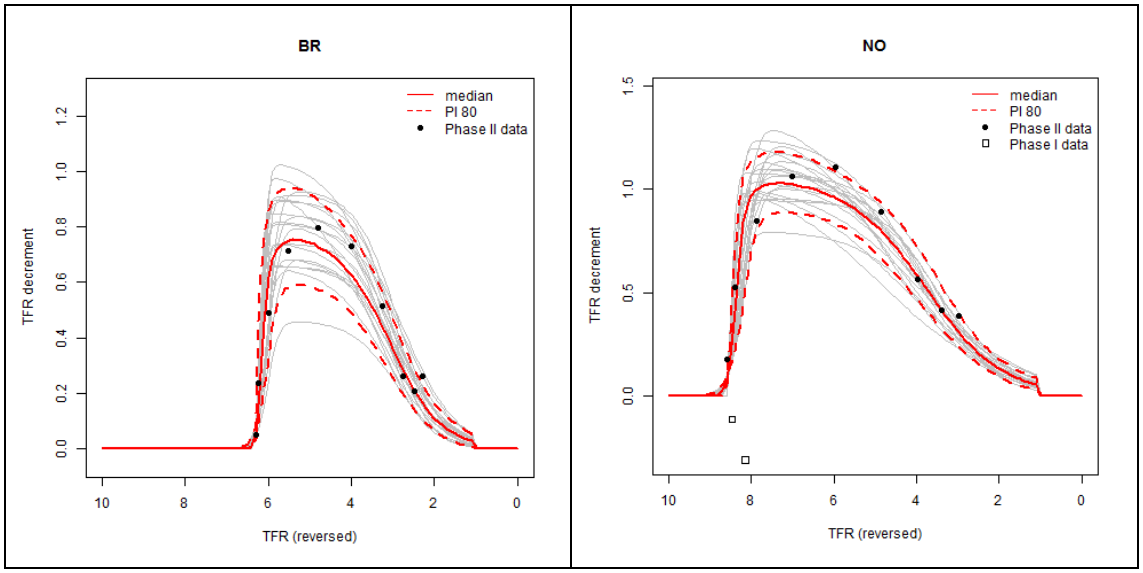
differences refers to the fertility levels when transition starts, with North Region presenting more than 8 children per women and South and Southeast around 6 children in the beginning of the transition.

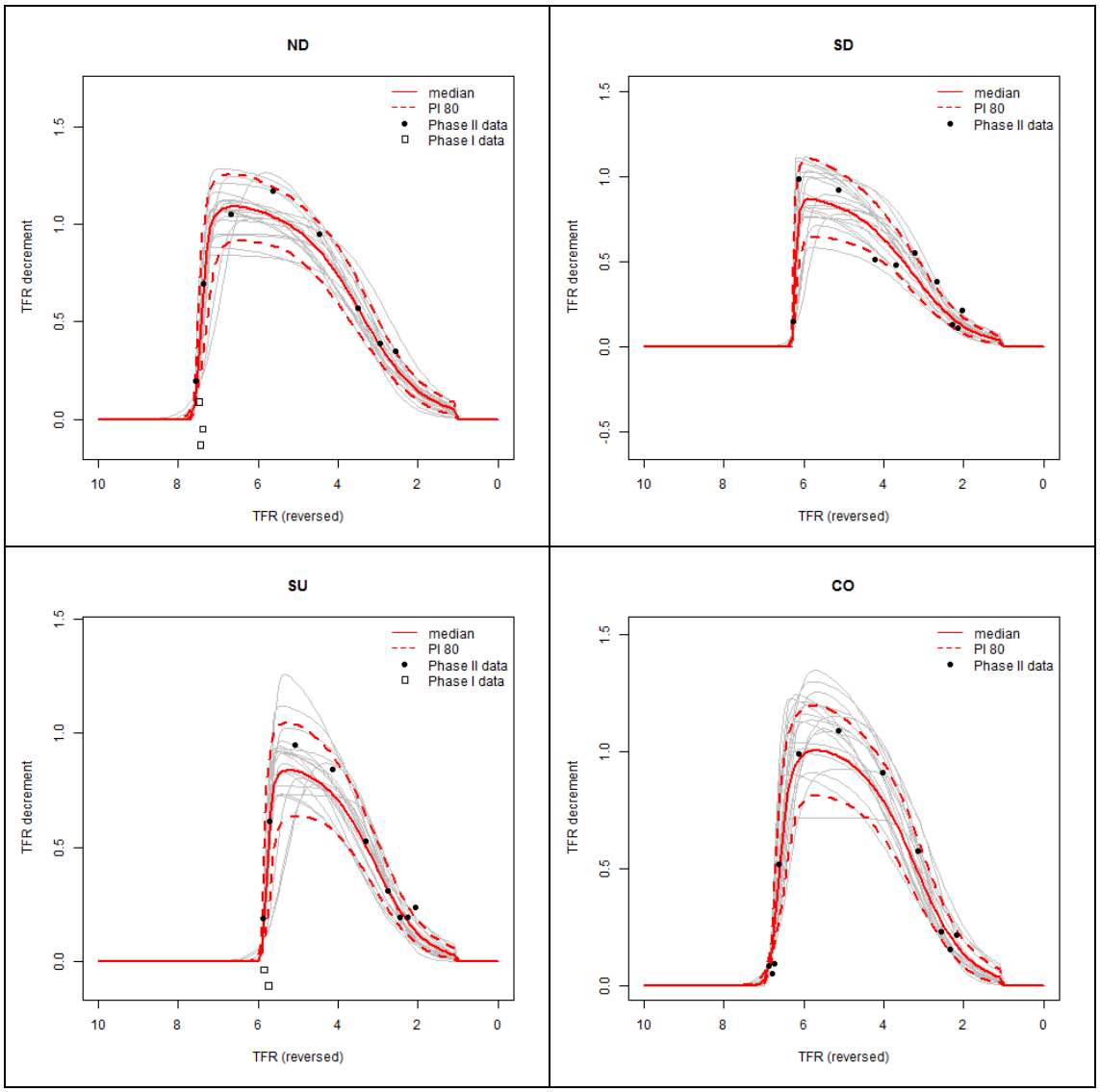
Another characteristic of the Brazilian transition is the velocity in which it has happened. The South and Southeast Regions presented a maximum decrement of 0.9 children per five-year period, while the other regions presented a decrement of 1.1. The Brazilian maximum decrement is lower than every region, since it represents a combination of different timings. Almost all UFs presented a maximum decrement of 1.0 children per five-year period or above. Fertility decrement of this magnitude has been observed only in a few countries in the world.

These singularities of the Brazilian fertility transition is one of the reasons that have made fertility predictions a difficult task in the country, with projections constantly overestimating the TFR.

Figure 4 also shows the recent acceleration of fertility decline for the country as a whole (BR) and for the Regions Southeast (SD), South (SU) and Central-West (CO). For Regions North (NO) and Northeast (ND) there was a reduction in the TFR decrement decline.

Figure 4 – Fertility Decline Curves – Brazil and Major Regions – 1950-2010





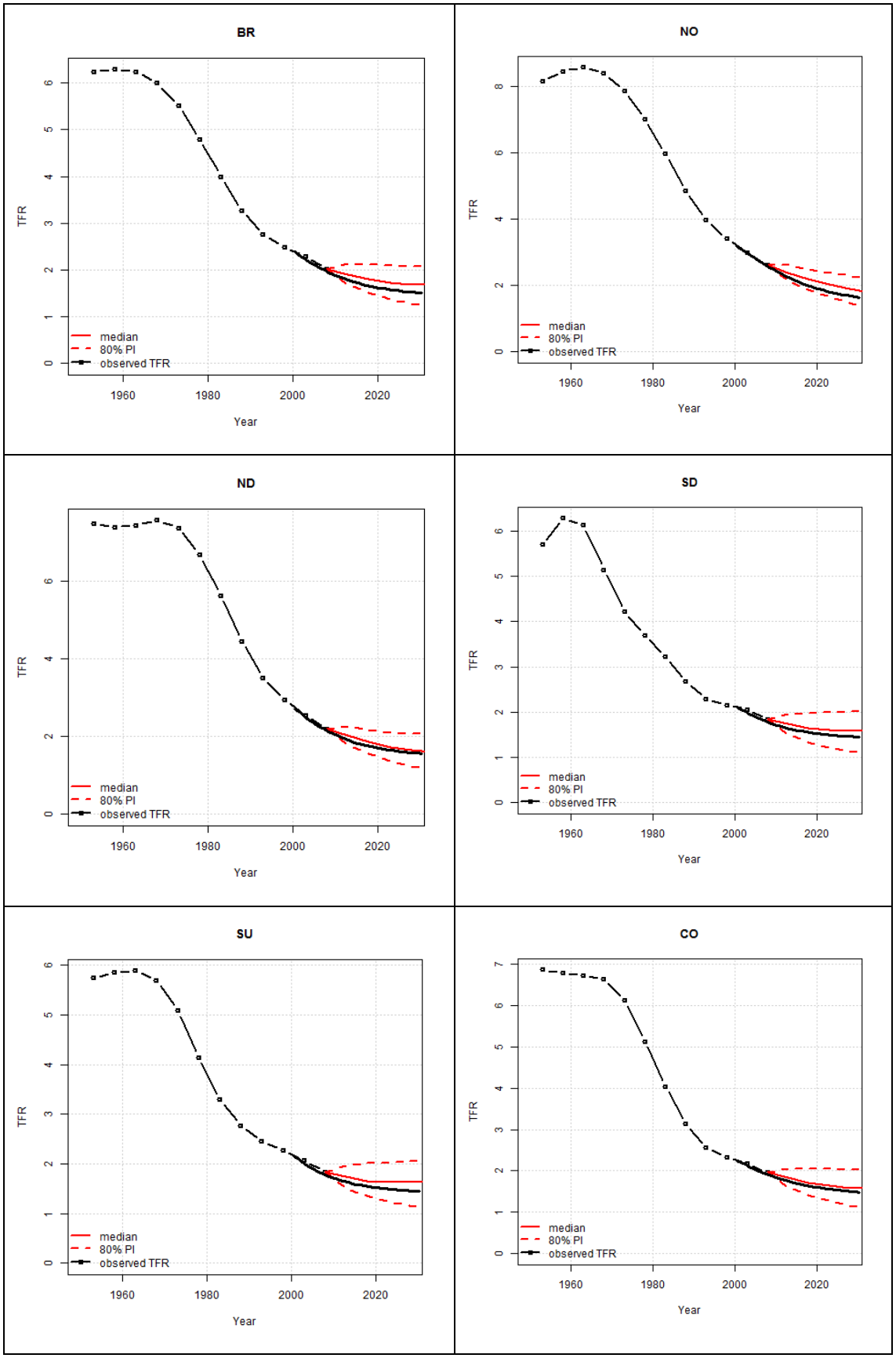
Source: (DeSA, 2013a; IBGE, 2013)

Figure 5 shows the point estimates and the 80% confident intervals, as well as the projections released by IBGE (2013). IBGE's projections are always lower than those projected using the Bayesian probabilistic approach. Central-West (CO) and Northeast (ND) Regions present similar estimations, while important differences arises in the North (NO). In the most developed regions, South (SU) and Southeast (SD), Bayesian estimation projects certain stability in fertility rates (around 1.6) during the 2020s, which precedes the future recovery, while IBGE projects decline achieving the limit of 1.45.

Figure 6 shows comparisons between TFR projected for the Units of Federation by the UN Bayesian method and by IBGE (2013) for every five years from 2015 to 2030. The main difference is found in Amapá (AP), the only state where IBGE's projections are out of the 80% confidence interval's range.

For 2015 and 2020, IBGE's estimations are systematically lower than those projected by the Bayesian model. For 2030, fertility projected by IBGE presented the regional convergence in different groups varying from 1.45 to 1.75, while projection using the Bayesian model presents a more heterogeneous distribution, ranging from 1.4 to 2.1 children per women.

Figure 5 – Total Fertility Rates estimated (1950-2010) and projected (2010-2030) by the Bayesian method (red lines) and by IBGE (black lines) - Brazil and Major Regions



Source: (DeSA, 2013a; IBGE, 2013)

Figure 6 – Total Fertility Rates projected by the UN Bayesian method and by IBGE - 2015-2030 – Units of Federation

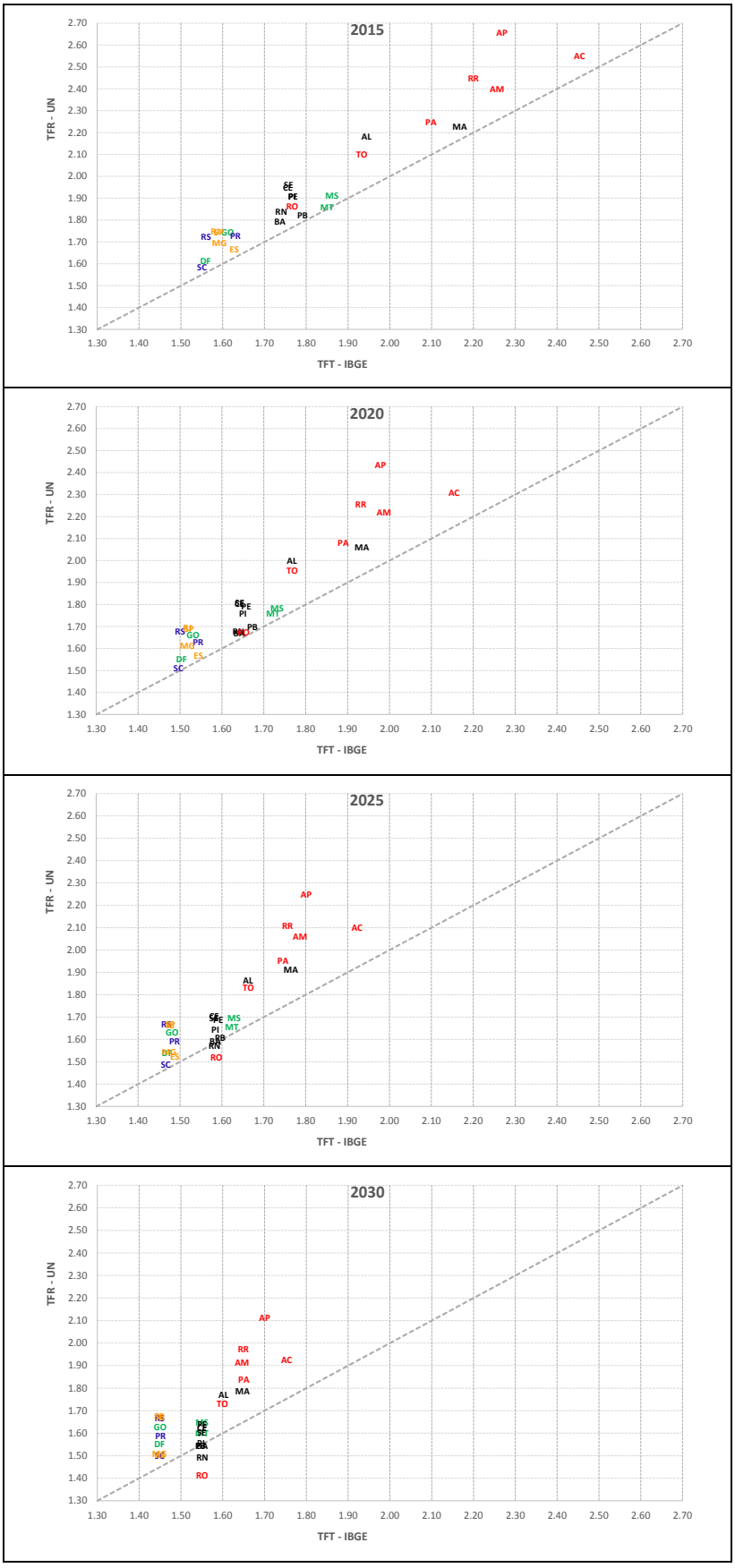
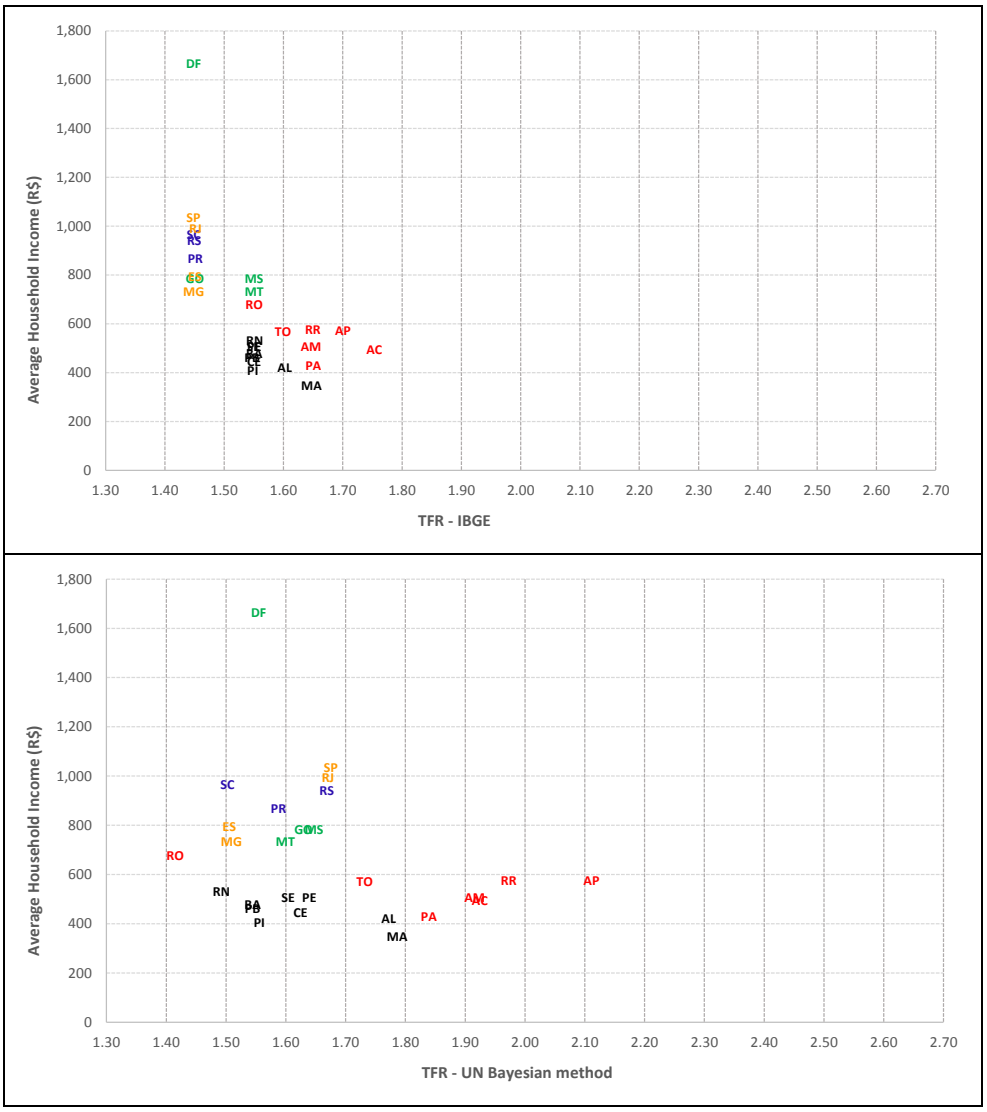


Figure 7 shows the relationship between the TFT projected for 2030 by each method and one of the covariates that are normally associated with fertility levels – average household income[5]. Assuming that there will still be some income inequalities in the country, and they will be related to the current income distribution, IBGE's projections show a clearer association between these two variables, while UN Bayesian projections leads to a more disperse relationship.

Figure 7 – Total Fertility Rates projected by the UN Bayesian method (2030) and by IBGE (2030) and average household income (in R$) in 2010 – Units of Federation



Source: (DeSA, 2013; IBGE, 2013; PNAD (2009;2011))

Discussion and Conclusions

We have shown that out-of sample projections systematically overestimate the observed TFR in 2010, since there was an acceleration of the fertility decline in the 2000s. Social advances observed in this decade might have influenced such rapid fertility decline.

For projections with starting point in 2010, IBGE's estimations still present lower figures than those projected by the Bayesian model, in addition to a different convergence pattern in 2030. Furthermore, for this year, IBGE's projections shows a clearer association

[5] Analyses using educational attainment variables show similar patterns.

between fertility levels and income and education. Those are extremely important variables to consider when projecting fertility. If improvements in income and education continue in the future, it would be expected the continuity of fertility decline in Brazil. The main question is if the extraordinary advances observed in the last decade will continue or not.

Another variable that must be included in fertility projections is the age-specific fertility rate. An additional particularity of Brazilian fertility, along with other Latin-American countries, is the high concentration of fertility among the young woman. Understanding the fertility schedule is central part of the relationship between income, education and fertility.

In addition to not consider uncertainty in the measures, another limitation of the method used by IBGE is suppose that fertility falls monotonically, ignoring that may be a recovery from below-replacement fertility. However, for the projection horizon used in this paper, only two UFs presented the end of fertility decline before 2030 using the Bayesian projection (Santa Catarina and Distrito Federal).

We can conclude saying that Bayesian probabilistic approach offers a promising alternative to sub-national level projections in Brazil, providing point estimations and, most importantly, measures of uncertainty. However, some adjustments might be added to the model such as the incorporation of age-specific fertility rates and some covariates related to fertility, such as educational attainment.

Furthermore, future works could try to use regions that have closer characteristics to the Brazilian context other than the level and the decrement of fertility decline in order to estimate the parameters. Latin-American countries, which have similar cultural and socioeconomic characteristics, in addition to a closer fertility age-structure, are potential candidates.

Bibliography

Abel, G.J., Bijak, J., and Raymer, J. (2010). A comparison of official population projections with Bayesian time series forecasts for England and Wales.

Alkema, L., Raftery, A.E., Gerland, P., Clark, S.J., Pelletier, F., Buettner, T., and Heilig, G.K. (2011). Probabilistic Projections of the Total Fertility Rate for All Countries. Demography *48*, 815–839.

Barros, R., Carvalho, M. de, Franco, S., and Mendonça, R. (2010). Determinantes da queda na desigualdade de renda no Brasil.

De Beer, J. (1992). Methods of fertility projections and forecasts.

Berquó, E., and Cavenaghi, S. (2014). Tendências dos diferenciais educacionais e econômicos da fecundidade no Brasil entre 2000 e 2010. XIX ENCONTRO Nac. Estud. Popul. An. São PedroSP Abep.

Booth, H. (2006). Demographic forecasting: 1980 to 2005 in review. Int. J. Forecast. *22*, 547–581.

Borges, G., and Silva, L. (2015). Fontes de dados de fecundidade no Brasil: características, vantagens e limitações. In Mudança Demográfica No Brasil No Início

Do Século XXI: Subsídios Para as Projeções de População Das Unidades Da Federação, (Rio de Janeiro: IBGE (in press)),.

Borges, G., Ervatti, L., and Silva, L. (2011). Desafios para o IBGE nas estimativas populacionais dos icípios brasileiros: aplicação de distintas metodologias. (Rio de Janeiro: ALAP/IBGE),.

Campos, M., and Borges, G. (2015). Projeção de níveis e padrões de fecundidade no Brasil. In Mudança Demográfica No Brasil No Início Do Século XXI: Subsídios Para as Projeções de População Das Unidades Da Federação, (Rio de Janeiro: IBGE (in press)),.

De Carvalho, J.A.M. (2004). Crescimento populacional e estrutura demográfica no Brasil (Cedeplar, Universidade Federal de Minas Gerais).

Cavenaghi, S., and Berquó, E. (2014). Perfil socioeconômico e demográfico da fecundidade no Brasil de 2000 a 2010.

CBED (1974). Projeção da população brasileira por idade e sexo - período 1970/2000. Rev. Bras. Estat. *35 (139)*, 257–370.

CEPAL, and CELADE (2004). América Latina y Caribe: Estimaciones y Proyecciones de Población 1950-2050.

DeSA, U. (2003). World population prospects: the 2002 revision (United Nations Publications).

DeSA, U. (2006). World population prospects: the 2004 revision (United Nations Publications).

DeSA, U. (2007). World population prospects: the 2006 revision. PLACE Popul. Div. Dep. Econ. Soc. Aff. UN Secr. Httpearthtrends Wri Orgtextpopulation-Heal.-379 Html.

DeSA, U. (2011). World Population Prospects: The 2010 Revision, Highlights and Advance Tables (Working Paper No ES/P/WP. 220. New York: United Nations, Department of Economic and Social Affairs, Population Division).

DeSA, U. (2013a). World population prospects: The 2012 revision. Popul. Div. Dep. Econ. Soc. Aff. U. N. Secr. N. Y.

DeSA, U. (2013b). World population prospects: The 2012 revision. Popul. Div. Dep. Econ. Soc. Aff. U. N. Secr. N. Y.

FIBGE, and CELADE (1984). Brasil: estimaciones y proyecciones de población, 1950-2025 (CELADE).

Gerland, P. (2011). Lessons learned and applications of the 2010 UN fertility projection methodology (Rio de Janeiro).

IBGE (2001). Projeção Preliminar da População do Brasil por Sexo e Idades Simples: 1980-2050 Revisão 2000 (Departamento de População e Indicadores Sociais. Divisão de Estudos e Análises da Dinâmica Demográfica.).

IBGE (2008). Projeção da população do Brasil por sexo e idade - 1980-2050 - Revisão 2008 (Rio de Janeiro: Coordenação de População e Indicadores Sociais - IBGE).

IBGE (2012). Censo Demográfico 2010 - Nupcialidade, fecundidade e migração: resultados da amostra. Rio Jan. IBGE.

IBGE (2013). Projeções da População: Brasil e Unidades da Federação (Rio de Janeiro: Coordenação de População e Indicadores Sociais - IBGE).

Lee, R.D. (1998). Probabilistic approaches to population forecasting. Popul. Dev. Rev. 156–190.

Lutz, W., and Goldstein, J.R. (2004). Introduction: How to deal with uncertainty in population forecasting? Int. Stat. Rev. *72*, 1–4.

MENDES, M. (1994). Projeção preliminar da população do Brasil para o período 1980-2020 Rio de Janeiro: IBGE, 1994. 21 p. Textos para discussão.

OLIVEIRA, J., and FERNANDES, F. (1996). Metodologia e Considerações acerca da Projeção de População do Brasil: 1980-2020. São Paulo Em Perspect. *10*, 116–123.

OLIVEIRA, J. de C., Albuquerque, F., and Lins, I.B. (2004). Projeção da população do Brasil por sexo e idade para o período 1980-2050–revisão 2004. Rio Jan. IBGE.

Raftery, A.E., Alkema, L., Gerland, P., Clark, S.J., Pelletier, F., Buettner, T., Heilig, G., Li, N., and Sevcikova, H. (2009). White paper: Probabilistic projections of the total fertility rate for all countries for the 2010 World Population Prospects.

Rios-Neto, E.L. (2000). Passado, presente e futuro da fecundidade: uma visão de idade, período e coorte. Rev. Bras. Estud. Popul. *17*, 5–15.

Ševcıková, H., and Gerland, P. (2013). Bayesian probabilistic population projections: do it yourself.

Ševcıková, H., Alkema, L., and Raftery, A.E. (2011). bayesTFR: An R package for probabilistic projections of the total fertility rate. J. Stat. Softw. *43*, 1–29.

Wilson, T., and Rees, P. (2005). Recent developments in population projection methodology: A review. Popul. Space Place *11*, 337–360.