Measuring the Timing and Pace of Fertility Decline in Brazil Using a Bayesian Spatial Estimation Procedure

Joseph E. Potter, Population Research Center, University of Texas at Austin

Renato M. Assunção, Department of Statistics, Federal University of Minas Gerais

Carl P. Schmertmann, Center for the Study of Population, Florida State University

Suzana M. Cavenaghi, ENCE, Instituto Brasileiro de Geografia e Estatística (IBGE)

E-mail: joe@prc.utexas.edu

Note: For presentation at the Annual Meeting of the Population Association of America, Los Angeles, March 30-April 1, 2006. Research for this paper was supported by a grant from NICHD (R01HD041528-01A1).

1. Introduction

Some of the primary conclusions from The European Fertility Project (EFP) were important to build a consensus view about both historical and contemporary fertility transitions – namely, that they are only weakly associated with changes in living conditions, modernization, or development. Since Coale's first summary (1973), the EFP results and their interpretation have faced serious scrutiny. In brief, it has been argued that 1) the EFP authors too quickly dismissed the possibility that fertility change was an adjustment to changing social and economic circumstances, and 2) that the failure of their models and data to explain either the onset of fertility decline or its pace resulted more from technical problems with the analysis than from any underlying reality about the nature of fertility transitions (Brown and Guinnane 2002). Technical problems included the large size of many of the provinces that constitute the units of analysis, the inadequacy of attempts to date the onset of the transition (Guinnane et al. 1994), the lack of appropriate indicators of economic circumstances, and the failure to account for endogeneity and unmeasured heterogeneity in statistical models. Another limitation of the EFP was that while the authors often provided a wealth of maps of fertility and its components, they did not utilize modern spatial statistics in their analyses.

Following on the EFP conclusions were a number of influential overviews of the experience of developing and developing countries that found a weak relationship between development indicators and the fertility transition (eg. Cleland and Wilson 1987; Bongaarts and Watkins 1996), as well as a spate of research aimed at assessing the role of diffusion processes and social interaction in the adoption of modern contraception and other practices related to fertility decline. With the passage of time, and the spread of fertility decline to ever larger fractions of the world's population, attention has shifted from the earlier to the later stages of the fertility transition. There is more interest now in the question of when fertility decline will stop than in when it will begin. Not only are there quite a few formerly developing countries where fertility has fallen to "lowest-low" levels, but there are countries where the fertility transition seems to have "stalled" at levels well above replacement (Westoff and Cross 2005; Bongaarts 2005).

A possible approach to shed light on both the EFP issues and those related to the end point of the fertility transition is to use data on the local experience for relatively small spatial units within countries. With large sample microdata, we are presently conducting a study of the Brazilian fertility transition to find out how levels and changes in social and economic conditions affected the timing and pace of fertility decline in this large, heterogeneous South American country. The microdata are from the five censuses carried out from 1960 to 2000, and have geographic identifiers as far down as the municipality (county) level. For some purposes, we have found it convenient to work with a set of 502 *microregions.* These are aggregations of municipalities that we built up so as to form comparable areas across the five points in time. Since the population of many of these units is quite small, even with our large samples (10% to 25%), conventional estimates of the TFR for the least populated these areas may be dominated by sampling noise, while TFR estimates for some other units will be based on the experience of hundreds of thousands of women.

On the positive side, we have a large number of spatial units, a great deal of variation between them in terms of both fertility and development, and a lot of change over the forty year period of observation. We are challenged, however, by two limitations.

First, estimates of fertility for some of the units are subject to considerable sampling variation, and, second, our time series for each unit is quite sparse, consisting of only five points. One way to overcome both of these limitations is to use spatial relationships to increase the signal-to-noise ratio and empirical support for the estimated schedules and trend patterns. As we have shown elsewhere (Assunção et al. 2005), the spatial correlation between neighboring areas can be taken advantage of to produce more stable estimates of fertility rates. But, in this instance, it can also help us to estimate trend models that capture the main features of local fertility transitions.

In this paper, we will present results from the application of a new Bayesian spatial method to estimate the timing and speed of the fertility transition in each of the 502 Brazilian microregions. The model we fit to the time series of TFR estimates for each microregion is a logistic curve. After presenting parameter estimates, we assess how well our estimated curves fit the data, and, in particular, where the logistic assumption seems to create large residuals. We then proceed to test some of the main hypotheses or stylized facts regarding the timing of the onset of the fertility transition, and the speed with which it proceeds. Finally, we discuss the implications of these results for forecasting the future of fertility in Brazil.

2. Data and Methods

The microdata used in this study are from a series of population censuses in Brazil covering a 40 year period (1960, 1970, 1991, and 2000). The information on fertility comes from long form questionnaires that were applied to a 25% sample in 1960, 1970, and 1980, and to around 12.5% of the population in 1991 and 2000. For the year preceding each census from 1970 on, we have the number of children born in the year preceding the census for each woman of reproductive age. We used this information to construct age-specific fertility rates for each microregion at each point in time. We then adjusted these rates based on the enumeration of children under age five, adjusted for an estimate of survivorship based on the census information on surviving and deceased children ever born. However, in 1960 the data on current fertility were not available, and the TFR for 1960 was estimated assuming that the age pattern of fertility was the same as in 1970 (Potter, Schmertmann, and Cavenaghi 2002).

The number of women in the sample, for each unit of analysis at any of the five points in time, varies dramatically (from 1,117 to 3,740,469). Also, the data for 1960 is not complete in terms of the coverage for the entire country. For historical administrative reasons, there are nine states for which there no microdata are available. Data is missing for all of the states in the North region, two states in the Northeast (Maranhão and Piauí), one in the Southeast (Espírito Santo), and one in the South (Santa Catarina).

To establish the comparable geographic regions used in our analysis, we began with a 1991 boundary file for municipalities. We then worked backwards and, later, forwards with documentation regarding the subdivision of municipalities (cartograms and lists of municipalities) for all other years. In 1960 Brazil had 2,766 municipalities, and by 2000 there were 5,507. For the longitudinal analysis, we constructed minimal municipal comparable areas, which by 2000 were a total of 2,672 geographical units. For this paper, we used a geographical unit aggregated at the next higher level, called a microregion. These geographical areas are defined by the Geographic Department of Brazilian Bureau of Census according to economic homogeneity and commercial transportation links. Based on the definition of microregions in 1991, we formed 502 areas that were comparable across the entire period and comprised of one or more comparable municipal areas. A more detailed description of how we constructed the boundary files at the microregion level is available in an earlier paper (Potter, Schmertmann and Cavenaghi 2002).

We use Bayesian spatial methods to study the timing and speed of the fertility transitions in each of the microregions by fitting a unique logistic time path for the TFR in the census years for which we have fertility data (1960, 1970, 1980, 1991, 2000). Let $F_i(t)$ be the TFR in region i and time t. Conditional on parameters yet to be specified, we assume that $F_i(t)$ is a random variable with normal distribution with variance $\sigma^2(t)$ and mean

$$E\left(F_i(t)\right) = Fpost_i + \left(Fpre_i - Fpost_i\right) / \left(1 + e^{-\beta_i (t-\tau_i)}\right)$$

In this model, $Fpre_i$ and $Fpost_i$ are the initial and final levels of TFR along a transition path; these two parameters are estimated from the data. Initially, we made the strong assumption that they were common to all regions, but later found that it was both desirable and feasible to relax that assumption so that values of each were estimated for each microregion. The parameter τ_i is the halfway point of the fertility transition in region *i* (i.e., the time at which transition is 50% complete), and β_i represents the speed of the transition in region *i.* Note that we expect $\beta_i < 0$, with larger absolute values representing more rapid fertility declines.

Because we assume that each microregion has a unique transition path ($\beta_i, \tau_i, Fpre_i, Fpost_i$), there are more than 2,000 parameters to estimate simultaneously for Brazil as a whole. If we assume independence of parameters across space, then two parameters have to be estimated separately in each area from only a handful of observations and estimates will be very noisy. It is therefore desirable to impose additional constraints. In the Bayesian approach this is accomplished by specifying a prior distribution in which neighboring microregions are likely to have similar parameter values. Specifically, we use a Markov random field (MRF) as a spatial prior for both β and τ, thereby building in an assumption that the map for each parameter is relatively smooth across space. For the fertility decline speed parameters (the betas), this MRF model implies that the distribution of β_i given the values of its neighboring values are normally distributed around $\overline{\beta}_i$, the arithmetic mean of parameter values among the n_i neighbors of area *i*. The inverse of the conditional distribution variance is given by $\rho_\beta n_i$, with ρ_β being called the precision parameter. Note that this precision parameter controls how similar *a priori* the β_i value is to its neighbors' mean value $\overline{\beta}_i$ and, in this sense, can be considered a measure of prior spatial correlation. We will have more to say about this precision parameter later.

This assumption effectively reduces the number of independent parameters to a number far below 2,000, allowing us to estimate time trend parameters for any given region by "borrowing strength" from its neighbors. By specifying MRF prior distributions for parameters ($\beta_1,...,\beta_N$), ($\tau_1,..., \tau_N$), ($Fpre_1,...,Fpre_N$), ($Fpost_1,...,Fpost_N$), in which neighboring areas are more likely to have similar parameter values, we make it likely that the parameters' posterior distributions (which are estimated from the combination of data and priors) will have the same property. The MRF prior distribution does not dictate the location of high and low parameters values on a map. Rather, it merely assumes that neighboring parameter values are dependent, and tend to be similar.

Bayesian inference proceeds by generating samples from the posterior distribution of all parameters. From these posterior samples we make inferences about parameters. For example, suppose that we simulate a large sample of S possible values for the speed of the transition in region i, $\{\beta_i^{(1)},...,\beta_i^{(S)}\}$, from the posterior distribution. We can then produce a point estimate for β_i using an estimate of the posterior distribution expected value (namely, the arithmetic mean of $\{\beta_i^{(1)},...,\beta_i^{(S)}\}$), or a credibility interval from the quantiles of $\{\beta_i^{(1)},...,\beta_i^{(S)}\}$.

We generate posterior distribution samples using Markov Chain Monte Carlo (MCMC) algorithms. Starting from arbitrary and valid parameter values (or *states*, in MCMC terminology), there are theorems proving that, under very general conditions (Smith and Roberts, 1993) if the chain runs long enough, a Markov chain state at step n will be distributed approximately according to the Markov chain stationary distribution. Then, the basic idea of MCMC algorithms is to design a Markov chain model which has the posterior distribution as its unique stationary distribution, start from arbitrary parameter values and, after a rather long burn-in period, start to collect the Markov chain states (or parameter values) generated. This produces a dependent sample rather than an independent sample but this does not prevent drawing inference by means of the ergodic theorem (Tierney, 1994).

We produced our estimates of the logistic curve parameters with WinBUGS 1.4, a well known and commonly used software package/language for MCMC estimation of Bayesian models (Spiegelhalter *et al.*, 2000).). Our priors include Markov Random Fields for all four parameters $(\beta_1,...,\beta_{502})$,…,$(Fpost_1,...,Fpost_{502})$. In the Markov fields, we further assume that the strength of spatial dependence across Brazil is uncertain, by modeling the variability of each parameter in a given area around its neighbors' mean value (the inverse of the precision parameter τ) as a hyperparameter with its own hyperprior. Using the β parameter to be specific, the precision parameter ρ_β has a prior distribution equal to a gamma distribution with parameters 0.01 and 0.01, which implies in mean and standard deviation equal to 1.0 and 10, respectively. All four precision parameters, ρ_β, ρ_τ, ρ_{post}, and ρ_{pre}, received this same prior distribution.

We centered the MRF prior distributions around global values which received prior distribution with large variance. For example, the vector ($\beta_1,...,\beta_{502}$) were centered a priori around a global value normally distributed with mean 0 and variance 100, allowing for a flat prior distribution over the range of likely values for this global parameter. We made similar choices for the other four spatially structured parameters. The variances of the $F_i(t)$ were modeled *a priori* as an inverse gamma with parameters 0.001 and 0.001.

All the WinBUGS runs were obtained with a burn-in of 100,000 runs followed by 200,000 additional runs from which we saved every 10-th value for the statistics. Therefore, each point estimate in the next section is based on a sample of 20,000 values. Usual tests of convergence and stability of results against widely different (but plausible) initial values were ensured.

One additional problem is the lack of observed values for $F_i(t)$ in some Northern areas when $t=1960$. To fit the models, we allow these missing data to be estimated as parameters by the Bayesian procedure. To do that, we used two alternatives: we gave initial values compatible with the successive observed values in each area as well as allowing the WinBUGS program to randomly generate them from the prior distribution. The results are virtually the same and we report only those from the first alternative.

3. Results

Preliminary Estimates and the Decision to Let Fpre and Fpost Vary

We first obtained preliminary estimates for fertility transitions over 1960-2000 in the 63 microregions in the state of São Paulo, and later for all 502 microregions (all of Brazil) assuming a single unique value for F_{pre} and F_{post}. For the 63 São Paulo transitions, the Bayesian point estimates of these levels were 6.26 and 1.55, respectively. When we then expanded the sample to all 502 microregions, these levels changed to substantially higher values: 7.14 and 1.98. This divergence, as well as examination of the residuals in the individual estimates motivated us to attempt to estimate models with more parameters, first letting F_{pre} vary, and then letting F_{post} vary as well. The results are presented in the next sub-section.

Estimates of Four Parameters for All 502 Microregions

We present our estimates, in the first instance, on the spatial grid by way of thematic maps. Figures 1 and 2 are maps showing the location in time of the transition in each microregion. The first map shows the year in which the transition was estimated to have reached 10 percent completion ($2.19725/\beta_i + \tau_i$), which, following convention, we will refer to as the "start" date. The second shows the year at which the transition in each microregion is estimated to have reached its halfway point (τ_i). The estimates of the *Start* dates Figure 1 show that the fertility transition started in selected areas of the South and Southeast regions well before 1960, the beginning of our period of observation, and that it eventually spread out to cover the rest of the country during the following thirty years. The estimates of the *Halfway* points in Figure 2 range from the mid and late 1960s in the "pioneer" microregions of the South and Southeast to 2002 when what appears to be the last transition was estimated to reach its midpoint in a remote municipality of the Amazon (Novo Olinda do Norte in the state of Amazonas).

The estimates of Beta shown in Figure 3 exhibit an inverse pattern with the slowest transitions found in the South and Southeast, and the fastest in the North and Northeast. Apparently, there was a tendency for transitions to be faster in the same regions where they were longest in coming. The estimates of Fpre and Fpost shown in Figures 4 and 5 present substantial regional differences in the range of the transitions, indicating that the fertility transitions would both begin and end at lower levels in the South and Southeast than in the North and Northeast.

To examine the influence of period on these transitions more directly, we have presented aspects of the distribution of Fpre, Fpost, and Beta (10^{th} percentile, median, and 90^{th} percentile) in tables in which the transitions are categorized according to the period in which they began (Table 1) or were centered (Table 2). These distributions while clearly showing substantial variation within a given range of start years or halfway-points, also show substantial shifts through time. Earlier transitions have, on average, lower levels of fertility at the onset of the transition, and later transitions are, on average, faster. The latter is especially pronounced in the classification by start date.

Goodness of Fit

We picked the logistic function to model the time series of TFRs in a microregion because of its correspondence with a stylized conception of the demographic transition, and because it offered a clear handle on both the timing and pace of the transition, as well as "forecasts" of its beginning and end points. But the real world may not be so well behaved, especially in the presence of economic cycles of boom and bust, hyperinflation, rapid cultural change, and shifts in the availability of methods of contraception and abortion. Thus, we while we hoped the model would fit the data reasonably well; we expected to find some deviations beyond those that might be expected to result from the non-neglible sampling variation present in the estimates for places with relatively small populations. Moreover, we hoped that this process of holding the data up to a standard might prove to be a useful means of identifying unusual patterns of fertility change.

Our first approach to assessing goodness of fit was to examine both absolute and relative differences between the data points (the estimated TFR derived from the census of the same year) and the point on the estimated logistic curve at the corresponding point in time. For this purpose, we ordered the deviations by size within the five major regions of the country, and displayed them in a graph for each census year. The absolute and relative deviations are shown in Figure 7, where there is a plot for each census year, and within each plot, the differences are ordered by region, and within region by the size and direction of the difference. This is a vast amount of information, and is indicative of the interplay between regions brought about by the Bayesian smoothing, sampling variation "noise", as well as deviations resulting from a mismatch between the logistic curve and reality. While we hope to have more to say about these deviations in later versions of this analysis, at the moment, we do not have an explanation for what is undoubtedly the most striking aspect of these plots: the comparatively small size of the deviations found in 1991.

The second thing we did was to examine the residuals for each microregion by way of individual plots showing the fitted and observed points, as well the estimated values of Fpre, Fpost, and the Halfway point of the transition. We have shown four such plots in Figure 8. The cases were selected by inspection as being representative of three different types of fit. The first plot is for the microregion containing the city of Porto Alegre, Rio

Grande do Sul, as well as some surrounding municipalities, and it shows that smoothing has pulled up the estimates for 1970 and especially 1960 to be above the very low levels observed in the census data. We see no reason to believe these fitted estimates, nor to trust the estimates of Fpre, or the estimated dates for the halfway point or start (1945) for this particular transition. While Porto Alegre provides an extreme example, many of the "pioneer" microregions selected for Figure 6 based on the observed level of fertility in 1960 also present problems with fit.

The second plot is for Aracatuba in the state of Sao Paulo. In this case, by the year 2000, the fitted curve has flattened out at a level that is about 11 percent higher than the observed rate, and there is an offsetting deviation in 1980 when the observed value of the TFR is above the fitted rate. Here fertility seems to be declining along a path that might lead to a floor around 1.5 or 1.6, and the logistic function plus smoothing provides what would seem to be a weak, probably biased grip on the far end of the transition.

The third selected microregion is Chapada dos Veadeiros in the Central West state of Goiás. This place had a very small population in 1960, and had less than 2,000 women of reproductive age in the 2000 census sample. Here, the deviations are substantial and in varying directions, but there is no reason to suspect that the model is not doing a good job of estimating reality. The fourth and last plot, for Brasileia in the state of Acre, is representative of the large majority of microregions in that it shows a good fit, and gives us no indication that the model does not correspond with the true path of fertility.

The Association of Timing and Pace of Decline with Development Indicators

Our primary motivation for obtaining estimates of parameters that would characterize the timing and speed of this set of "local" fertility transitions was to see how well they corresponded to some of the main hypotheses or stylized facts regarding fertility transitions. Here the main issue has been whether the transition is best viewed as: 1) an adaptation to changing social and economic conditions, as the original theorists of the demographic transition would have it; or, alternatively, 2) controlled fertility mostly represents a behavioral innovation that spreads by way of social interactions through different regions and strata of the population, in large part, independently of social and economic conditions. As of the present moment, the consensus is that the second interpretation more closely corresponds with the historical evidence available from the countries that were included in the European Fertility Project, as well as the large mass of information that has been assembled on the fertility transitions that have taken place in developing countries since WW II.

One of the most cited and persuasive overviews of these questions (Bongaarts and Watkins 1996) included an analysis of the experience of 69 developing countries between 1960 and 1990. Their questions and analysis would seem to apply with equal or greater force to the transitions experienced by local areas within countries. Thus, as a starting point for our own analysis, we have selected several of their main findings to see if they are consistent with the transitions we observe and assess in Brazilian microregions. The first of these is that development threshold for the onset of fertility fall through time, presumably as a result of the influence of the example of other countries in the same region that have already begun the transition in fertility. The question in our case is whether the level of development at the start of the transition fell through time as the fertility decline spread out from a minority of places in the South and Southeast of Brazil to the rest of the country.

To address this question, we chose four indicators: the average number of years of education attained by women 15-49 in a microregion; the proportion of persons in the labor force whose occupation was in the primary sector; the proportion of households located in urban areas; and the proportion of households which had electricity. For transitions beginning after 1960, we estimated the level of the corresponding indicator at the start year by way of a linear interpolation of the values registered in the respective census years. The distribution of these indicators for the transitions starting in successive five year periods is shown in Table 3. Here, there appears to be a difference between the indicators for the transitions beginning in the first period, 1960-1964, and those for the later transitions. The indicators at the start of the transitions in this first group show a slightly higher level of educational attainment, a smaller primary sector, and noticeably more urbanization and electrification. However, from 1965 onwards, no consistent trend is apparent. There is a gradual increase in the average levels of education and electrification, but, on the other hand, there is an increase in the proportion of people employed in the primary sector and a decrease in urbanization.

An intriguing question that we are not yet in a position to answer fully concerns the level of development of the 62 microregions that began their transitions before 1960. Looking at their indicators for 1960 (not shown) they clearly were better off than those in which the onset occurred afterwards. However, we will need to piece together what we can with respect to the relevant indicators from published tabulations from the 1940 and 1950 censuses to get a better grip on the levels at the estimated time of onset. Indeed, it will also be interesting to see how much we can learn from the published data regarding the level of the TFR in these earlier censuses since we have reason to suspect that our estimates of both the start date and Fpre may be quite unreliable for these "pioneers". Parenthetically, to provide a picture of just where this early low fertility was located we prepared a map (Figure 6) showing the level of fertility and location of all the microregions for which we had data with a TFR below 4.5 in 1960.

If we shift our focus from the onset point to looking at "how far along" in their transitions local areas were at different levels of development, and then comparing across censuses, a secular shift becomes apparent. For this purpose, we chose the two indicators that our earlier analysis (Potter, Schmertmann, and Cavenaghi 2002) had identified as being most closely associated with the level of fertility in a fixed-effects model. In Table 4, we show the median level of "completeness" of the transition among microregions cross-classified according both education and electrification in the respective censuses. At any joint level of development, the median transition level shifts upwards as one moves from earlier to more recent censuses. Of course, some of this shift is due to a shifting composition within cells, but certainly not all of it.

The second major finding in the Bongaarts and Watkins (1996) study that we would like to address with our Brazilian data concerns the relationship between the pace of the fertility transition and development. Their conclusion is that the pace of the transition is unrelated to the speed with which development indicators change during the transition. However, they find that pace is related to the level of development at the start of transition, as well as time elapsed since another country in the same region entered the transition. We have argued that our own results from estimating a fixed-effects model (Potter, Schmertmann, and Cavenaghi 2002) contradict the first of these conclusions. While we could examine the association between our Beta parameter and change in development indicators, we suspect that doing so would not add much to the previous analysis.

One issue that we can address here, however, is the finding that pace is related to the level of development at the time of onset as well as time elapsed since other places in the region entered the transition. To that end, we ran a simple, unweighted OLS regression with Beta as the dependent variable, and our estimates of the level of development at the time of onset as well as the date of the halfway point of the transition as the predictors. The estimated coefficients and significance levels are shown in Table 5. These results show that the level of education at onset was inversely associated with the speed of the transition, while a later halfway point and higher electrification at onset were positively associated with pace. The estimated effect of the timing of the transition is consistent with our earlier examination of the shifts in our parameters through time (Table 2), as well as with the results for developing country transitions. However, the evidence that the level of development at onset influences the pace of decline seems to differ from the Bongaarts and Watkins (1996) results.

4. Discussion

As the annual meeting approaches, we have not yet fully digested the results presented here. Nevertheless, we are pleased with them. The logistic model seems to fit the bulk (about 80%) of the 502 transitions quite well, and provides a clear picture of when and at what TFR the transitions began, when and at what TFR they will end, and how fast they were. There is a lot to indicate that local transitions in Brazil have not and will not stall at levels far above replacement. Similarly, in most cases, it seems that fertility well below replacement is not an immediate concern.

The analysis has also served to draw our attention toward a minority of cases where the model does not seem to yield reliable or fully believable results. These cases are of two types. The first group consists of the pioneer microregions in which fertility was already quite low in 1960. It seems that in these places, there may have been a relatively long and slow decline that took place over a length of time that we cannot know or do not yet know. It seems to us that unlikely that such slow then faster transitions can or should be forced into the logistic mold, and that they are worthy of more attention than they have received so far. The second group consists of those places where there are strong indications that the transition is not going to level off at the estimated value of Fpost, and the TFR has either reached or seems likely to reach a value below 1.8. Again, these transitions seem to warrant further attention once we develop a good means of identifying them. In particular, it would be to know if there is anything about their socioeconomic trajectory that is different from the others.

Apart from the adequacy of the logistic model, both of these outlying groups show the weaknesses as well as the strengths of the Bayesian methodology that we have adopted. We use it because it helps us with both sampling variation, and the limited number of observations for each individual transition, as well as the large time interval between observations. "Borrowing strength from neighbors" seems to work most but certainly not all of the time. We need to spend some time thinking about how to further evaluate and explain the magic that WinBUGS has wrought.

With regard to the long-standing but now quite subdued debate concerning innovation versus adjustment perspectives on the fertility transition, we believe our modeling and analysis has something to add. The results presented here would seem to fall in between the extreme positions. There is no doubt that later transitions are faster, and the existence and example of the earlier transitions must have had some influence on the later transitions. However, the threshold for the onset of the fertility transition seems to have been remarkably stable, at least after 1965. Indeed, the fact that we have identified a relatively large number of places with low fertility in 1960 that may have had low fertility coupled with high development indicators for what may have been quite a long period raises an interesting question. Why did it take so long for others to follow their example and pick up on this innovative behavior? Further work with dusty census volumes as well as the microdata will have to be done, but perhaps the experience of the pioneer local regions will pose a fairly severe challenge to the innovation perspective.

Finally, there are clearly a number of ways that this type of analysis could be extended. One which we have long had in mind but have not yet implemented is to incorporate a logistic curve for mortality, and tie it to the curve for fertility. In addition to incorporating mortality, we envision two modeling approaches that could overcome the disjuncture in our current analysis between “smoothed” estimates of the fertility parameters, and “raw” estimates of the development covariates. In the first, after obtaining a point estimate, the posterior mean, for each parameter $(\beta_1,\ldots,\beta_N)$, $(\tau_1,\ldots,\tau_N)$, we could then use these estimates as inputs to a second stage, in which we analyze how a microregion’s transition path (β_i,τ_i) is related to its socioeconomic conditions.

Another possibility is to introduce a hierarchical structure in the Bayesian framework allowing the parameters to be both spatially correlated and dependent on covariate values. Hence, the covariates affect the demographic processes through the parameters controlling the timing and speed of the transition. A general model would be similar to one which puts

$$\beta_i = \gamma_0 + \gamma_1 X_{i1} + \ldots + \gamma_k X_{ik} + \varepsilon_i$$

where ε_i follows a spatial autoregressive or similar distribution. Possibly, different covariate subsets could explain different parameters. That is, the covariates explaining the speed transition do not have to be the same as those explaining the transition timing, especially in terms of their location in time, and levels versus differences. Proceeding in this way, would lead to a highly structured model, with several hierarchical levels, but with interpretable parameters with possibly great explanatory power.

REFERENCES

- Assunção, Renato, Carl P. Schmertmann, Joseph E. Potter, and Suzana M. Cavenaghi. 2005. Empirical Bayes Estimation of Demographic Schedules for Small Areas. *Demography* **42**(3): 537-558.
- Bongaarts, J. 2005. The causes of stalling fertility transitions in the developing world. Presented at the XXV International Population Conference, IUSSP, Tours, France.
- Bongaarts, John; Watkins, Susan C. 1996. Social interactions and contemporary fertility transitions. *Population and Development Review*, Vol. 22, No. 4, Dec. pp. 639-82.
- Brown, J. C. and T. W. Guinnane. 2002. Fertility transition in a rural, Catholic population: Bavaria, 1880-1910. *Population Studies* **56** (1): 35-49.
- Coale, A.J. 1973. "The Demographic Transition Reconsidered." Presented at International Union for the Scientific Study of Population, Liège, Belgium.
- Cleland, J. and C. Wilson. 1987. Demand theories of the fertility transition: An iconoclastic view. *Population Studies* 41:5-30.
- Guinnane, T.W., B.S. Okun, and J. Trussell. 1994. "What Do We Know About the Timing of Fertility Transitions in Europe." *Demography* 31: 1-20.
- Potter, Joseph E., Carl P. Schmertmann, and Suzana M. Cavenaghi. 2002. Fertility and Development: Evidence from Brazil. *Demography* **39**(4): 739-761.
- Roberts, G. O., and Smith, A. F. M. (1993) Simple conditions for the convergence of the Gibbs sampler and Metropolis-Hastings algorithms. *Stochastic Processes and Their Applications*, 49, 207-216.
- Spiegelhalter, D.J., Thomas, A., Best, N., and Gilks, W. (2000) *WinBUGS: Baysian Inference Using Gibbs Sampling. Version 1.3*. MRC Biostatistics Unit, Cambridge UK. (BUGS homepage http://www.mrc-bsu.cam.ac.uk/bugs/)
- Tierney, L. (1994) Markov Chains for exploring the posterior distributions (with discussion). *Annals of Statistics*, 22, 1701-1762.
- Westoff, C. F. and Cross, A. R. 2005. The Kenya stall. Presented at the XXV International Population Conference, IUSSP, Tours, France.

Figure 1 – Estimates of Start Year

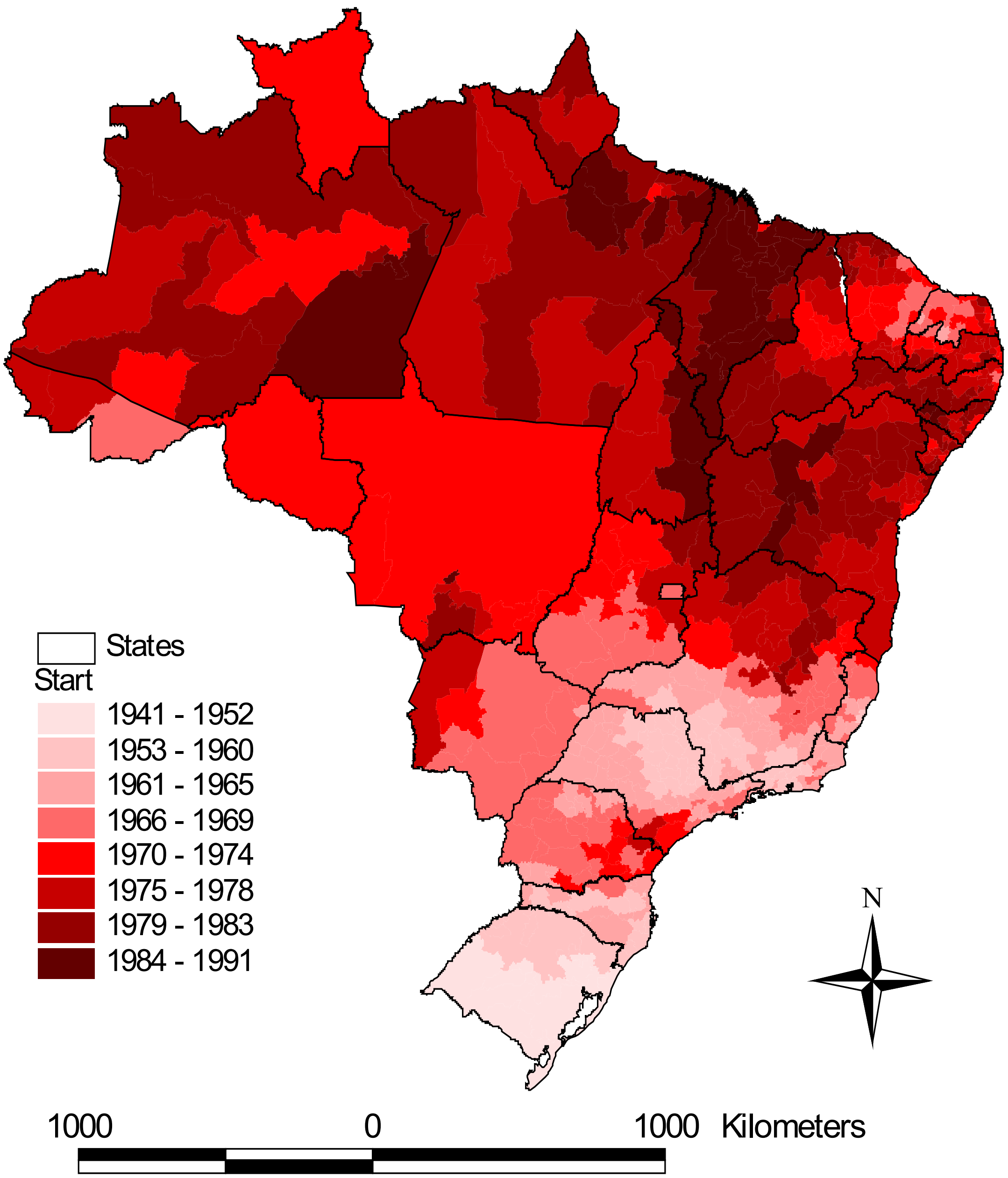


Figure 2 – Estimates of Halfway Point

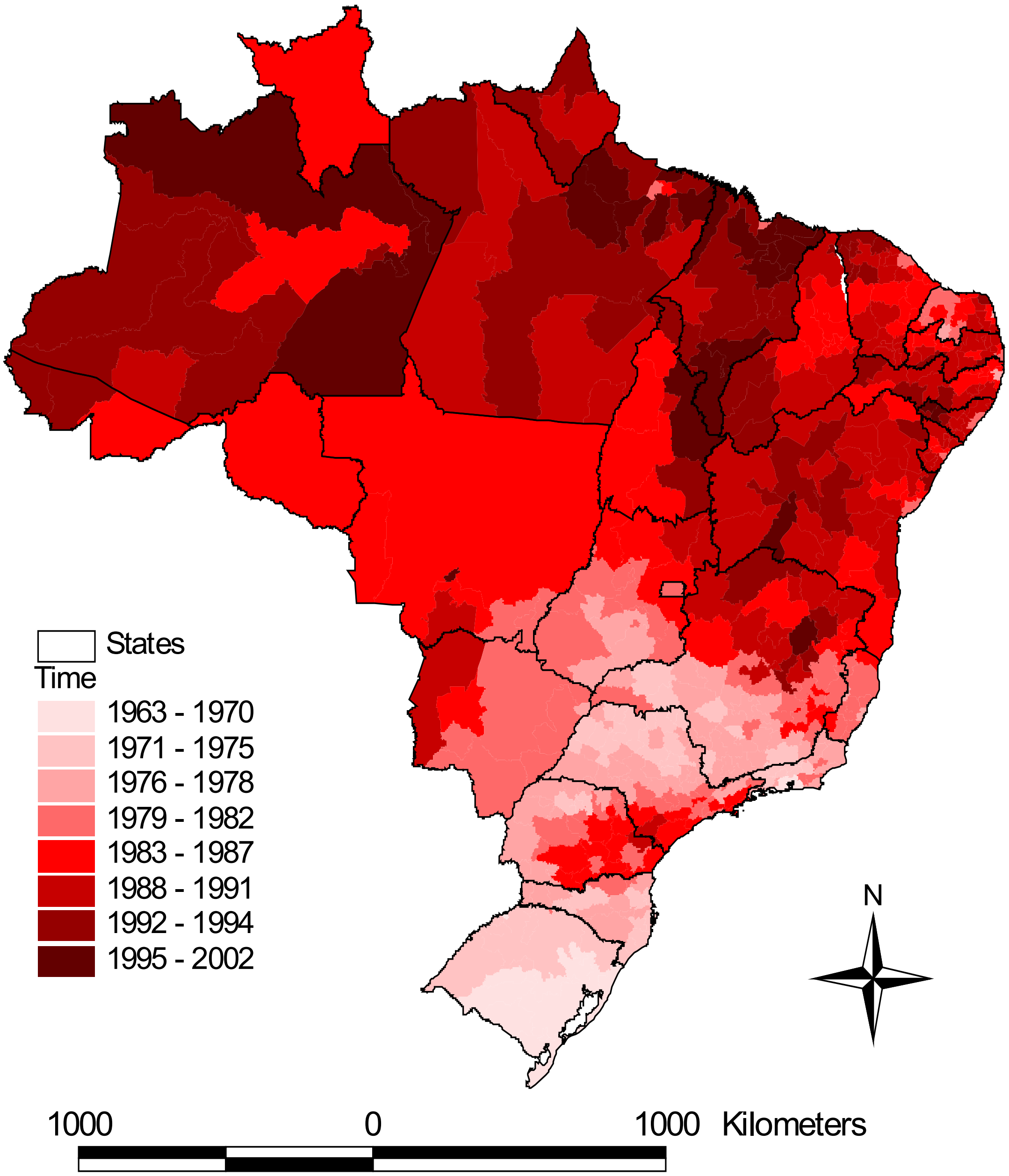


Figure 3 – Estimates of Beta

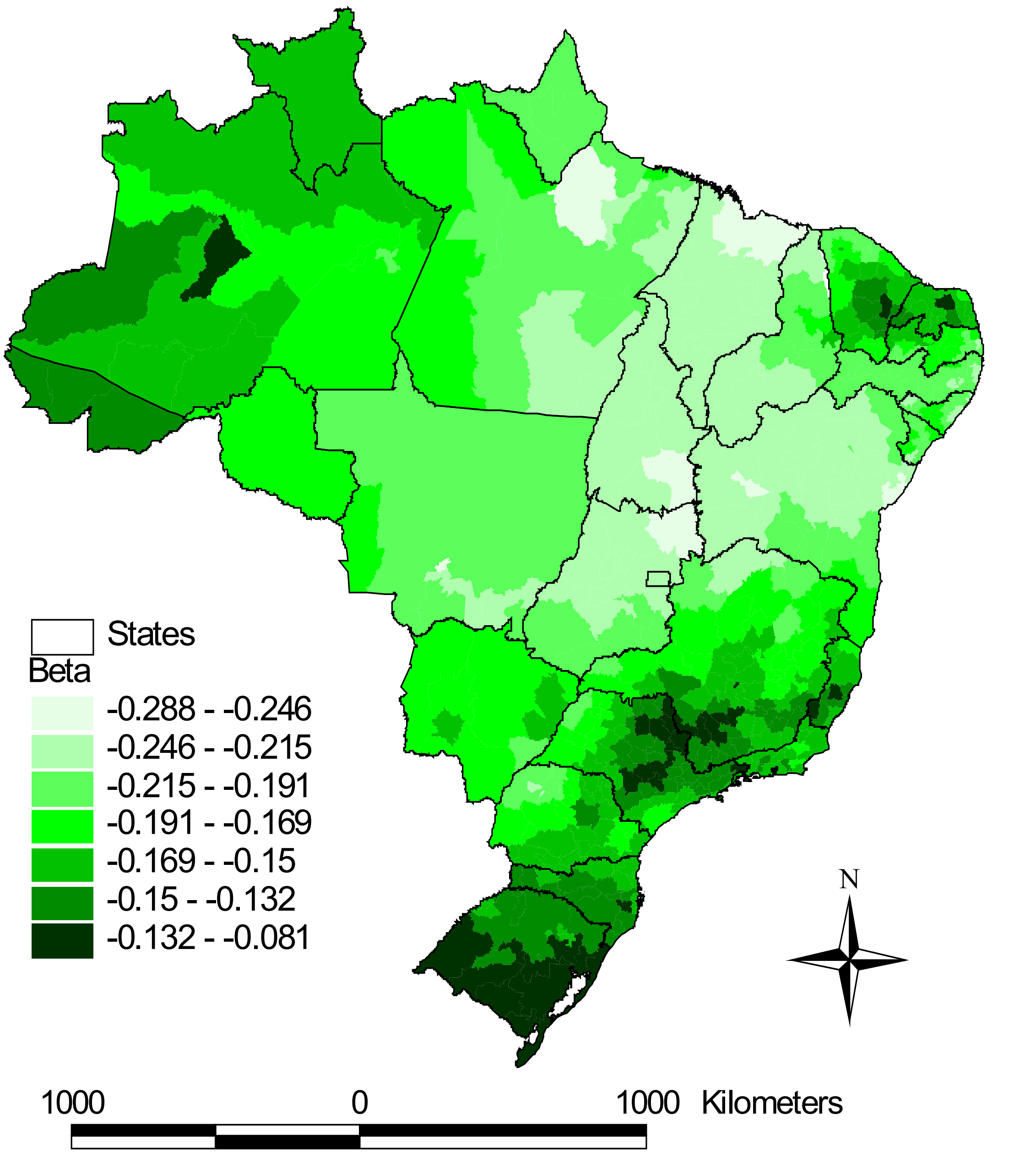


Figure 4 – Estimates of Fpre

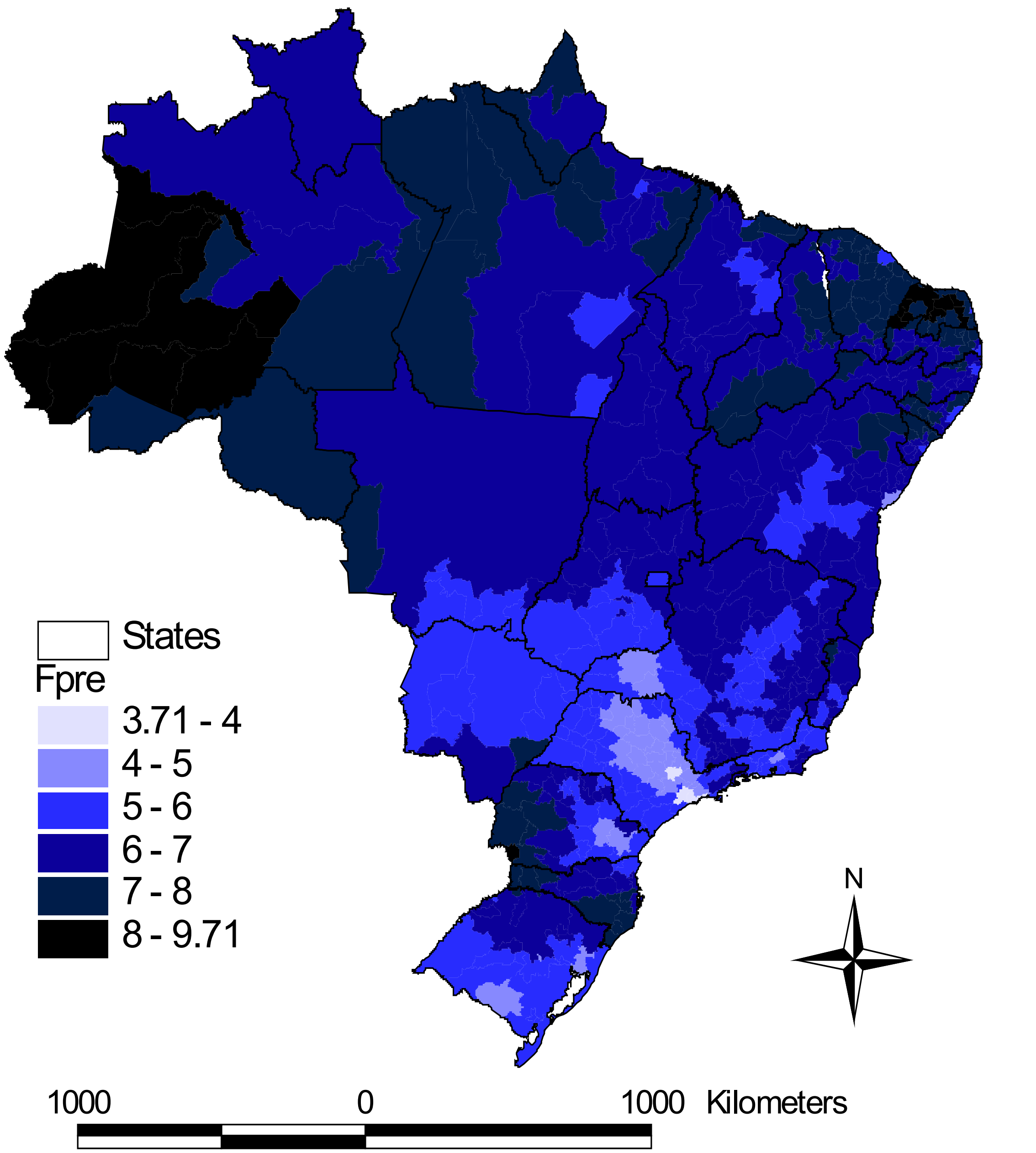


Figure 5 – Estimates of Fpost

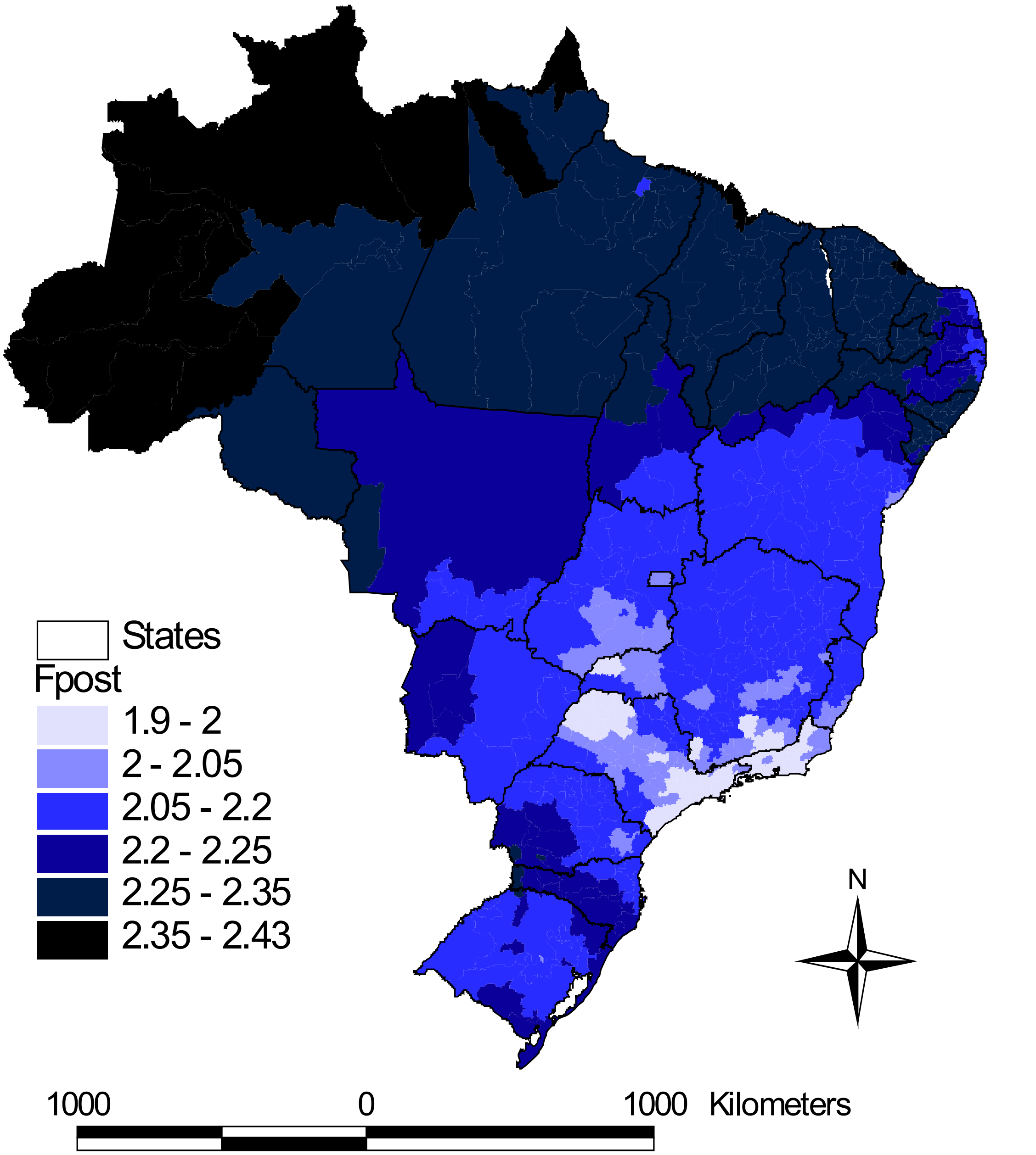
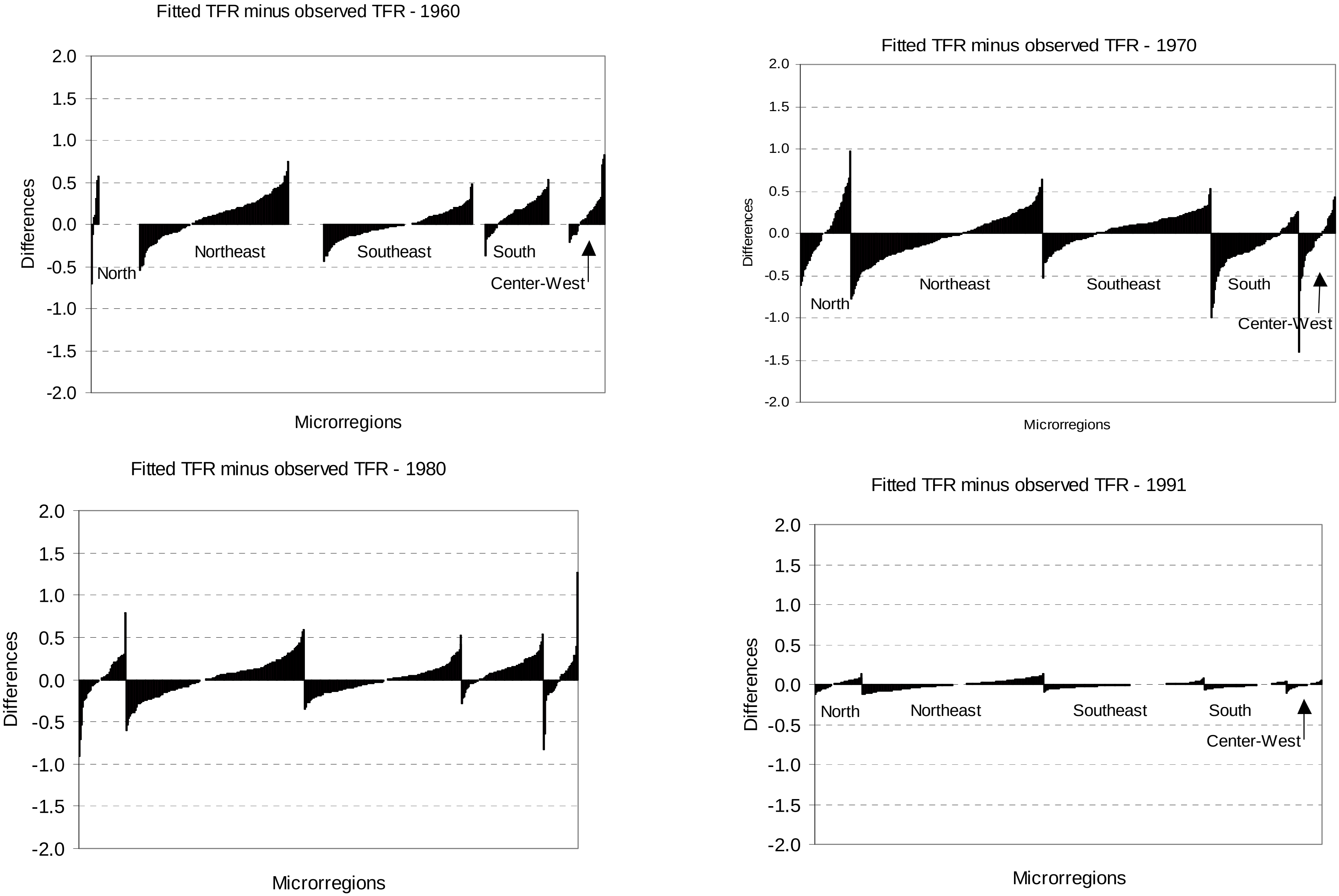
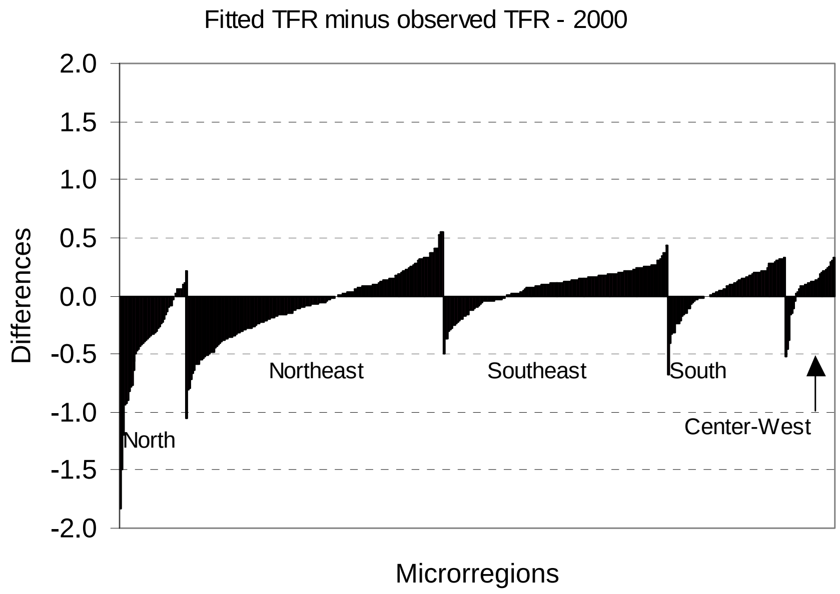
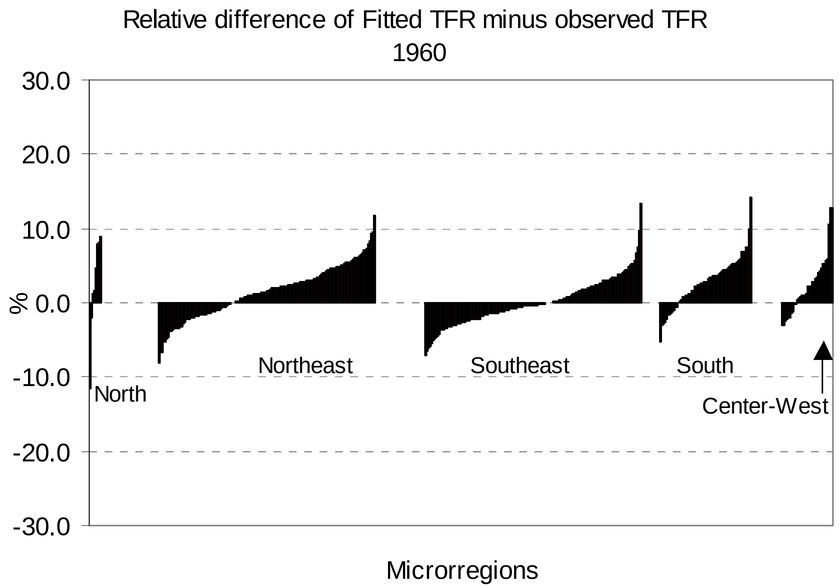


Figure 6 -- Places with TFR less than 4.5 in 1960 (unadjusted Census estimates).

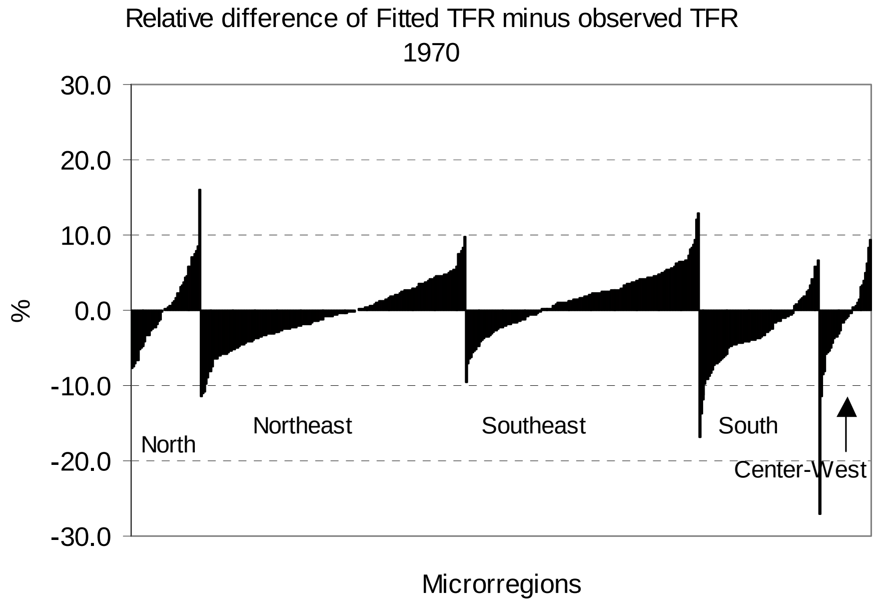
Figure 7. Plots of Absolute and Relative Differences, by Region and Census Year



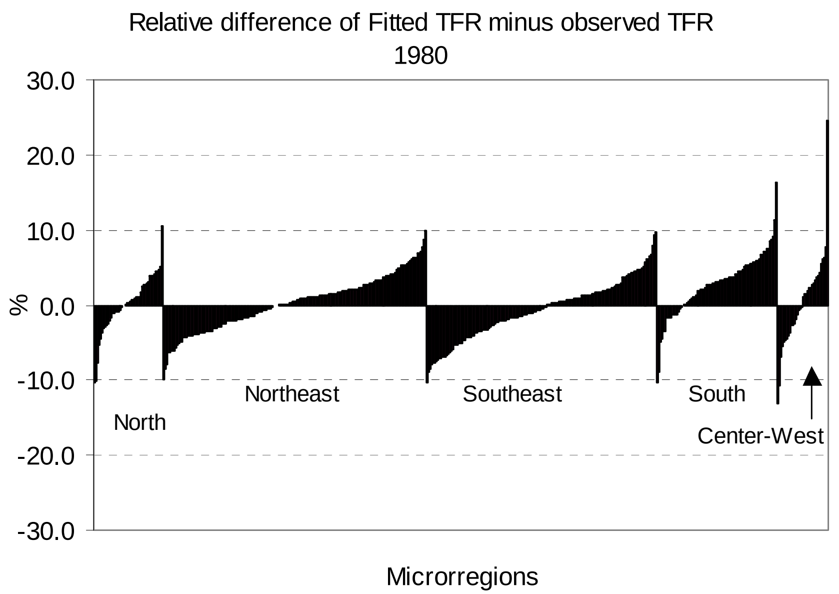

Fitted TFR minus observed TFR - 2000
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
Differences
North
Northeast
Southeast
South
Center-West
Microrregions



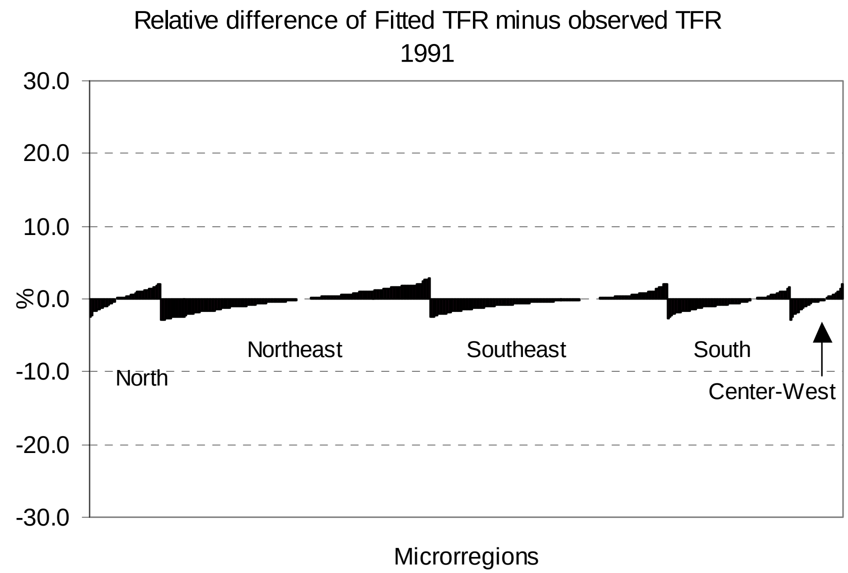
Relative difference of Fitted TFR minus observed TFR
1960
30.0
20.0
10.0
0.0
-10.0
-20.0
-30.0
%
North
Northeast
Southeast
South
Center-West
Microrregions



Relative difference of Fitted TFR minus observed TFR
1970
30.0
20.0
10.0
0.0
-10.0
-20.0
-30.0
%
North
Northeast
Southeast
South
Center-West
Microrregions



Relative difference of Fitted TFR minus observed TFR
1980
30.0
20.0
10.0
0.0
-10.0
-20.0
-30.0
%
North
Northeast
Southeast
South
Center-West
Microrregions



Relative difference of Fitted TFR minus observed TFR
1991
30.0
20.0
10.0
0.0
-10.0
-20.0
-30.0
%
North
Northeast
Southeast
South
Center-West
Microrregions

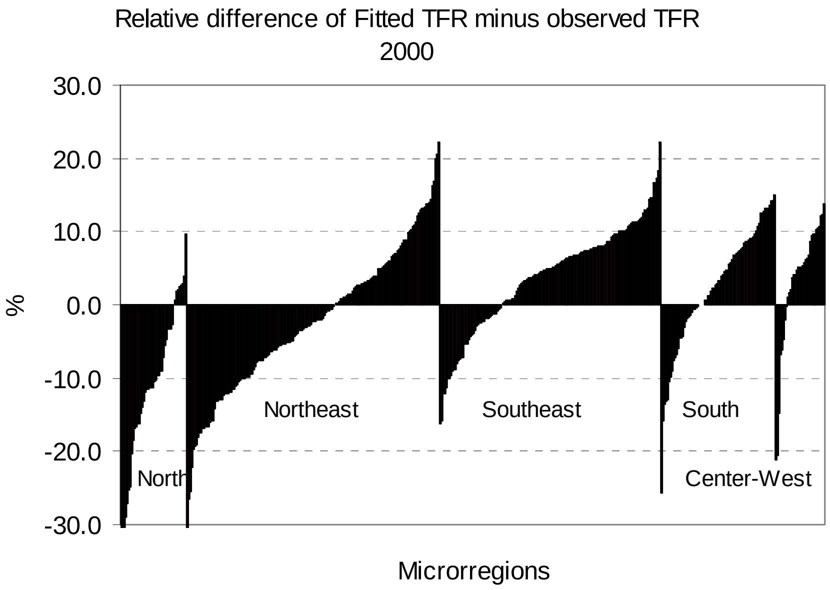
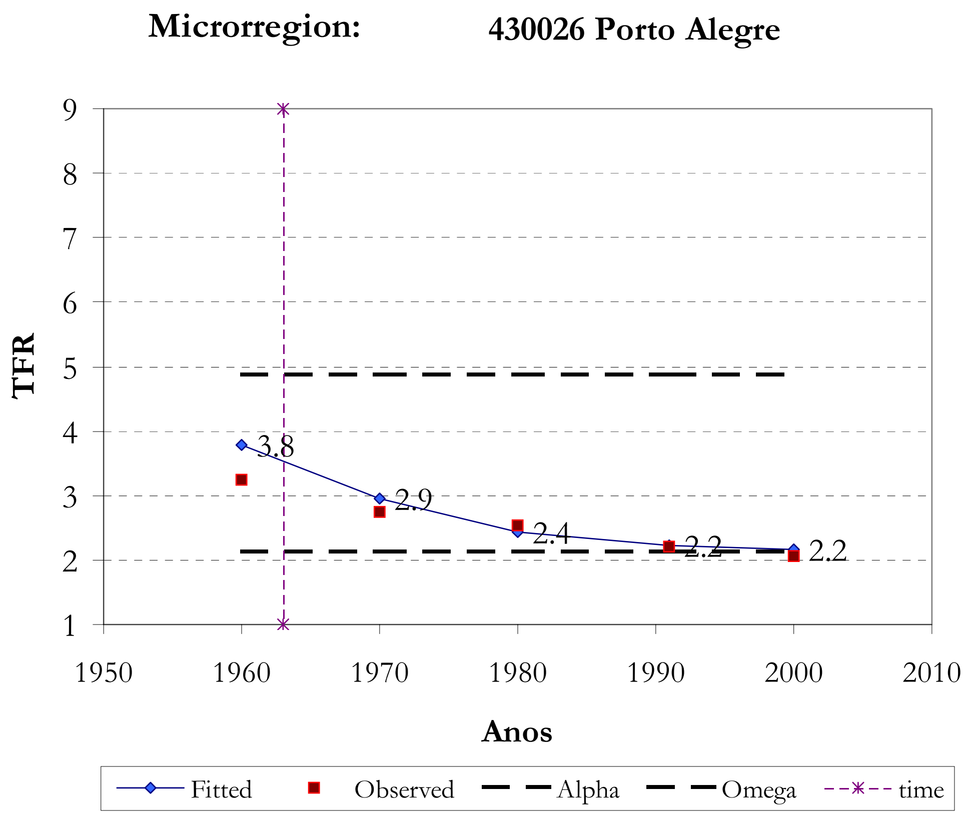
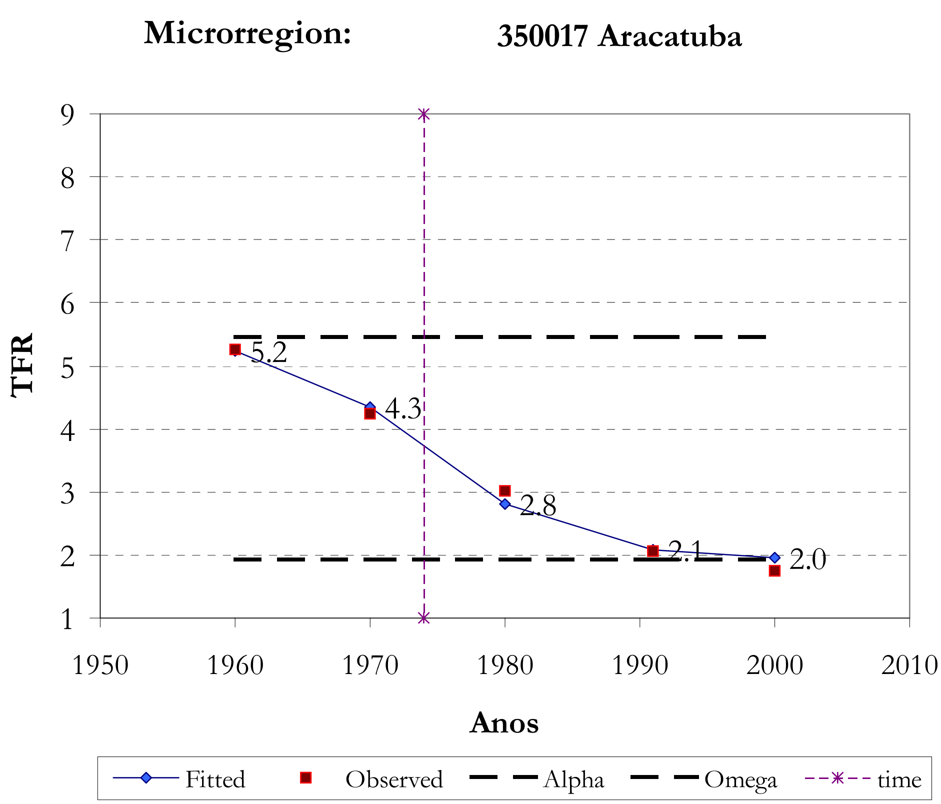
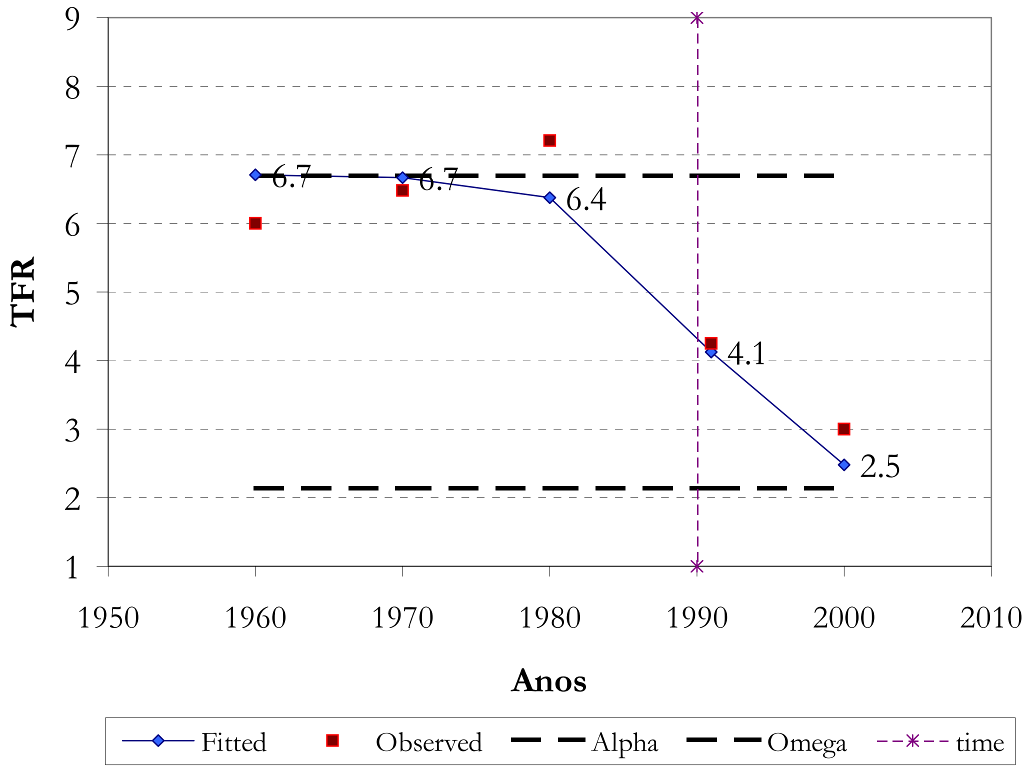

Relative difference of Fitted TFR minus observed TFR
2000
30.0
20.0
10.0
0.0
-10.0
-20.0
-30.0
%
North
Northeast
Southeast
South
Center-West
Microrregions

Figure 8. Plots for Selected Microregions





Microrregion: 520005 Chapada dos Veadeiros



Microrregion: 120004 Brasileia

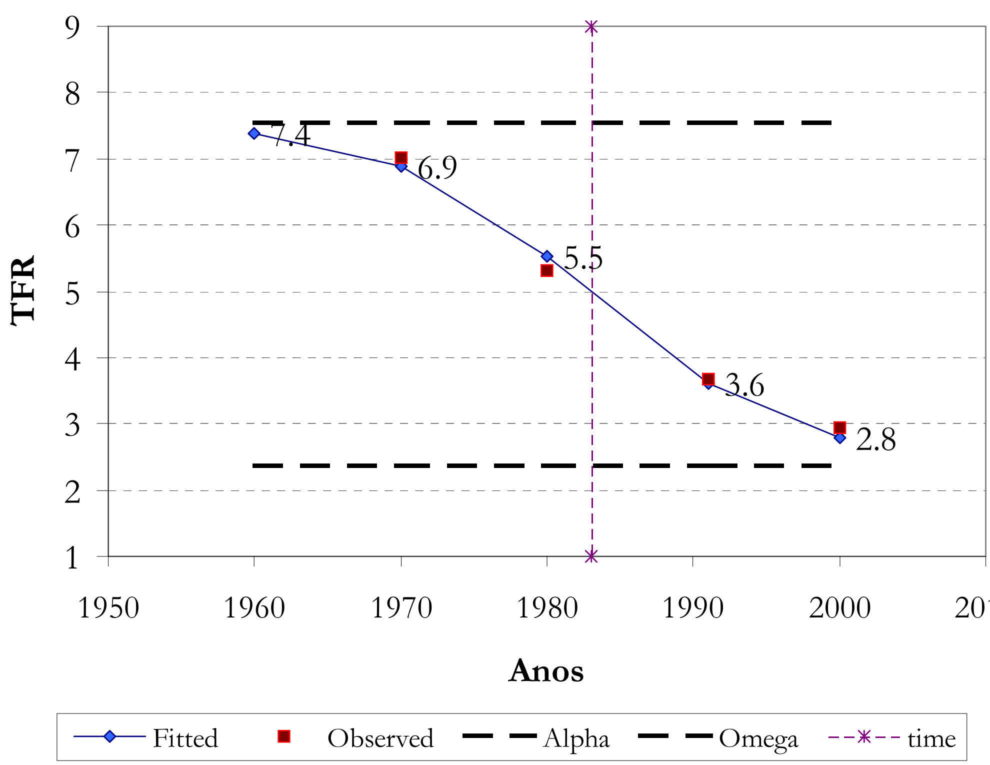


Table 1. Distribution of Estimates of Fpre, Fpost, and Beta, by Estimated Start Year

Start	F_{pre}			F_{post}			Beta		
	10%	50%	90%	10%	50%	90%	10%	50%	90%
1950 - 1954 (n=17)	4.93	5.55	6.27	2.06	2.20	2.23	-0.14	-0.12	-0.10
1955 - 1959 (=45)	4.56	5.35	6.70	1.99	2.07	2.21	-0.15	-0.13	-0.12
1960 - 1964 (n=90)	4.80	5.83	7.18	1.97	2.04	2.21	-0.18	-0.15	-0.13
1965 - 1969 (n=87)	5.22	6.34	7.56	1.98	2.10	2.31	-0.20	-0.16	-0.14
1970 - 1974 (n=66)	5.46	6.60	7.65	2.07	2.21	2.34	-0.23	-0.17	-0.15
1975 - 1979 (n=106)	6.04	6.76	7.60	2.10	2.26	2.33	-0.22	-0.20	-0.16
1980 - 1984 (n=66)	6.00	6.70	7.47	2.13	2.26	2.34	-0.23	-0.20	-0.17
1985 + (n=25)	6.04	6.67	7.60	2.24	2.31	2.34	-0.27	-0.24	-0.20

Table 2. Distribution of Estimates of Fpre, Fpost, and Beta, by Estimated Year of Halfway Point

Halfway Year	F_{pre}			F_{post}			Beta		
	10%	50%	90%	10%	50%	90%	10%	50%	90%
1965 - 1969 (n=14)	4.93	5.47	5.93	2.06	2.18	2.23	-0.14	-0.12	-0.11
1970 - 1974 (=32)	4.60	5.64	6.94	1.96	2.05	2.22	-0.17	-0.15	-0.13
1975 - 1979 (n=129)	4.70	5.78	7.18	2.00	2.06	2.23	-0.20	-0.15	-0.13
1980 - 1984 (n=95)	5.08	6.30	7.48	2.02	2.13	2.33	-0.21	-0.16	-0.13
1985 - 1989 (n=105)	5.73	6.78	7.60	2.09	2.24	2.33	-0.23	-0.18	-0.15
1990 - 1994 (n=100)	6.00	6.78	7.66	2.13	2.28	2.35	-0.23	-0.20	-0.16
1995 + (n=27)	5.97	6.78	7.64	2.17	2.32	2.35	-0.27	-0.23	-0.18

Table 3. Distribution of Estimates Development Indicators, by Estimated Start Year

Start	Women's Education			Primary Sector			Urban			Electrification		
	10%	50%	90%	10%	50%	90%	10%	50%	90%	10%	50%	90%
1960 - 1964 (n=73)	1.83	2.20	3.03	23.95	60.10	72.80	0.33	0.44	0.73	23.61	39.22	68.20
1965 - 1979 (n=78)	1.30	1.91	3.05	19.80	69.91	84.88	0.21	0.36	0.74	7.50	21.24	58.70
1971 - 1974 (=66)	1.23	1.92	3.44	22.86	66.15	84.35	0.22	0.36	0.74	10.00	19.10	55.73
1975 - 1979 (n=106)	1.34	2.00	2.87	48.63	67.45	81.25	0.23	0.37	0.55	14.75	24.25	39.60
1980 - 1984 (n=66)	1.45	2.23	2.95	59.62	70.88	79.85	0.19	0.33	0.47	16.93	26.63	38.65
1985 + (n=25)	1.95	2.54	3.26	61.66	73.34	81.56	2.00	0.30	0.44	23.30	29.62	39.13

Table 4. Median Transition Level among Microregions, by Level of Education and Electrification

Mean Educ	% Electrification				
1960	(0,20]	(20,40]	(40,60]	(60,80]	(80,100]
(0,2]	**1**	**4**	**7**	**-**	**-**
(2,4]	**7**	**10**	**8**	**10**	**8**
(4,6]	**-**	**-**	**-**	**-**	**41**
(6,8]	**-**	**-**	**-**	**-**	**-**
(8,10]	**-**	**-**	**-**	**-**	**-**
1970					
(0,2]	**2**	**4**	**4**	**-**	**-**
(2,4]	**8**	**23**	**28**	**31**	**27**
(4,6]	**-**	**4**	**60**	**28**	**32**
(6,8]	**-**	**-**	**-**	**-**	**-**
(8,10]	**-**	**-**	**-**	**-**	**-**
1980					
(0,2]	**8**	**11**	**-**	**-**	**-**
(2,4]	**8**	**19**	**38**	**50**	**36**
(4,6]	**-**	**-**	**61**	**63**	**66**
(6,8]	**-**	**-**	`	**61**	**81**
(8,10]	**-**	**-**	**-**	**-**	**-**
1991					
(0,2]	**-**	**39**	**42**	**-**	**-**
(2,4]	**-**	**37**	**50**	**59**	**54**
(4,6]	**-**	**-**	**58**	**75**	**88**
(6,8]	**-**	**-**	**-**	**-**	**93**
(8,10]	**-**	**-**	**-**	**-**	**-**
2000					
(0,2]	**-**	**-**	**-**	**-**	**-**
(2,4]	**-**	**-**	**75**	**76**	**86**
(4,6]	**-**	**-**	**87**	**85**	**90**
(6,8]	**-**	**-**	**-**	**87**	**97**
(8,10]	**-**	**-**	**-**	**-**	**98**

Table 5. OLS Model for Beta regressed on indicators of development at onset and the halfway point date

Beta	Coefficient	P > |T|
Women's Education	-0.01820	0.000
Primary Sector	0.00005	0.795
Urban	-0.00676	0.746
Electrification	0.00063	0.000
Halfway Point	-0.00219	0.000
Constant	4.18370	0.000
R^2	0.281	