Published in final edited form as:

Popul Stud (Camb). 2013 November ; 67(3): . doi:10.1080/00324728.2013.795602.

Bayes plus Brass: Estimating Total Fertility for Many Small Areas from Sparse Census Data

Carl P. Schmertmann,
Florida State University

Suzana M. Cavenaghi,
Escola Nacional de Ciências Estatísticas

Renato M. Assunção, and
Universidade Federal de Minas Gerais

Joseph E. Potter
University of Texas-Austin

Abstract

Small-area fertility estimates are valuable for analysing demographic change, and important for local planning and population projection. In countries lacking complete vital registration, however, small-area estimates are possible only from sparse survey or census data that are potentially unreliable. Such estimation requires new methods for old problems: procedures must be automated if thousands of estimates are required, they must deal with extreme sampling variability in many areas, and they should also incorporate corrections for possible data errors. We present a two-step algorithm for estimating total fertility in such circumstances, and we illustrate by applying the method to 2000 Brazilian Census data for over five thousand municipalities. Our proposed algorithm first smoothes local age-specific rates using Empirical Bayes methods, and then applies a new variant of Brass's P/F parity correction procedure that is robust under conditions of rapid fertility decline.

Keywords

fertility; indirect estimation; spatial statistics; Bayesian statistics; Empirical Bayes; Brazil

Introduction

In the 1960s and 1970s, indirect estimation of demographic parameters for populations with incomplete vital registration was one of the crown jewels of population science. Indirect techniques were based on demographic theory, and used ingenious combinations of data and models that were likely to be robust against common errors. The authors of *The Demography of Tropical Africa* (Brass et al. 1968) and *UN Manual IV* (Coale and Demeny 1967) pioneered this work. A second generation of demographers, most of whom have now celebrated their 60th birthdays, developed it further. Indirect techniques for demographic estimation were among the first tasks addressed by the US National Academy of Science's Committee on Population and Demography when it was established in 1977, ultimately leading to the publication of *UN Manual X* (United Nations 1983). They were the main staple of the demography taught and later extended at the UN Regional Demographic Centres, and were the subject of scores of journal articles and applied reports. For Brazil, the

country that serves as an example in this paper, the most prominent application of indirect methods was Carvalho's (1974) reconstruction of demographic histories for large regions.

The golden age of indirect estimation came to an end with the widespread implementation of large, nationally representative surveys in developing countries. The World Fertility Survey and its successors provided the impetus and the financing for much of this data collection. These internationally sponsored surveys were often complemented by similar surveys conducted by national statistical agencies, population councils, or ministries of health. Survey data proved to be reasonably reliable not only for estimating fertility, but also infant and child mortality. What indirect estimation remained was largely confined, in most countries, to the analysis of adult mortality.

In this paper, we suggest that there may now be new uses for indirect methods in small-area estimation. Even the largest national surveys rarely have sufficient sample sizes to produce reliable direct estimates of rates for small geographic areas. If registration of births and deaths remains incomplete, demographers may face a situation very similar to the one that prevailed in the 1960s and early 1970s. The only alternative is to take advantage of information from both vital registration and censuses, knowing that such information may be defective or incomplete.

Demand for reliable demographic estimates for small areas seems to be growing rapidly in many countries, irrespective of the adequacy of vital registration. This increased interest probably stems from efforts to decentralise and target social policies, and from the importance of demographic parameters and population sizes to such efforts. The interest of international organizations in developing local indicators (e.g., the UNDP's Human Development Index) has also increased the demand for small-area estimates.

Local-level estimation with incomplete data, however, poses two additional problems that were absent from earlier applications to large populations. First, even with census data and vital registration, the number of recorded events and the populations exposed to risk may be too small to permit precise estimation. This problem may be especially acute in countries, like Brazil, that have reached low fertility levels while vital registration is still incomplete. The second new difficulty derives from the sheer number of estimates. Most methods proposed to correct data and to estimate fertility or mortality indirectly depend on the researcher's judgment and interpretation for the selection of the best estimates. Making literally thousands of such judgments would be a cumbersome task whose results would be difficult to reproduce.

Estimation Problem and Strategy

We use a concrete example to illustrate the challenges of small-area estimation, and to explain our proposed solutions. Consider the problem of simultaneously estimating period total fertility (TFR) for 5506 small Brazilian administrative regions, known as *municípios*, for the year prior to the 2000 census. Census long-form questionnaires are the only adequate data source for this task, as we elaborate in the next section. However, even with aggregation into standard five-year age groups and a very large Census sample (more than 5.4 million women aged 15–49, reporting more than 380 thousand births in the year before the census), many of the $5506 \times 7 = 38542$ (municipality, age) cells have very few women or births from which to estimate period fertility rates.

Table 1 provides information about the range of available census sample sizes. The first seven rows present distributions within five-year age groups. The last row presents the distribution of total municipal sample sizes, with women of all ages combined. The majority of municipal samples have fewer than 500 sampled women, and fewer than 40 births in the

last year, from which to estimate total fertility. Problems of small samples for rate estimation are even more acute within some age groups, particularly for births to women 30+ years old. In short, estimating age-specific fertility rates directly from census data (local births in the last year/local women) will produce very noisy and unreliable estimates for many locations. Many estimated local rates (22%, or precisely 8519 of 38452) will be zero. In addition, for many cells with only a few sampled women, small random fluctuations in local births will produce implausibly high or low rate estimates.

Figure 1 provides a specific example, to which we will return several times over the course of this article. The data are for the least populous Brazilian municipality in 2000: Borá, in the state of São Paulo. Borá's 2000 census sample consisted of only $n=34$ women 15–49, who reported six births in the previous year. Direct estimates of age-specific fertility rates yield values above .30 for age groups 15–19 and 25–29, and zeroes for the last three age groups. The direct estimate of Borá's total fertility (TFR=5.37) is extremely high when compared to Brazil's national fertility pattern in 2000, and it is even more implausible because fertility in municipalities near Borá (a neighbourhood that we define more precisely later) was below the national average in all age groups.

Small-sample problems like these are of course familiar to demographers everywhere. In the Brazilian case, estimation difficulties are compounded by possible reporting errors common in poor country surveys (reference period mistakes, underreporting of children who were born out of wedlock or have died, and so forth). Further, rapid declines in fertility levels over the years preceding the census in many parts of Brazil (Potter et al. 2002) complicate the use of standard correction methods, like P/F ratios, that assume recent stability in fertility levels.

We propose a two-part strategy to address these difficulties. First, use Empirical Bayes (EB) estimation of local fertility and parity schedules (Assunção et al. 2005) to “borrow strength” from data in neighbouring areas. This improves the signal-to-noise ratio in local census estimates such as those in Figure 1. Second, generalise the standard P/F correction procedure to allow for changing fertility patterns over the reproductive lifetimes of the women in the sample. One can then apply the new P/F adjustment procedure to EB estimates of total fertility and parity for each municipality, in order to produce a final set of estimates that we call *EBB (Empirical Bayes plus Brass)*. We describe the census data and our EBB estimation strategy in detail in the next sections.

Brazilian Fertility Data

Birth records are available in Brazil from a vital registration system, the *Registro Civil*, organized by the Bureau of the Census (IBGE). Data come from offices of Notaries Public, which should theoretically register all births. Coverage has improved notably in the last decade, but there are still high rates of under-registration in poor regions, especially in northern and north-eastern Brazil. For example, IBGE (2008) estimates that 35% of all births in 1998 were not registered within the mandatory 90-day period for the *Registro Civil*, and that state-level rates of under-registration in 1998 varied from 10% in São Paulo to 81% in Maranhão. For 2008, IBGE estimates that national under-registration was much lower, at 10% (ranging from 2% in São Paulo to 37% in Amazonas state).

Because vital registration is incomplete, almost all published estimates of Brazilian fertility rates, at all levels of geography, use data from Censuses and surveys. The main source for our estimates in this article, microdata files from IBGE's 2000 Demographic Census long-form questionnaire, include information for women aged 15–49 about the number of children ever born and about the number of integer years since the last live birth.

The sampling fraction for the 2000 Census long form was either 10 or 20 per cent (for municipalities with an estimated population larger or smaller than 15,000 inhabitants, respectively). As illustrated above in Table 1, most municipal sample sizes are quite modest, and some are extremely small.

The 2000 Census microdata have person-level weights to account for the two different sampling fractions, and for non-response. Calculations in this paper use these weights, but control the totals to the original sample size. In other words, our calculations and sample sizes refer to weighted, but not expanded, sample data from the Census.

Stage 1: Empirical Bayes Smoothing of P and F schedules

Small local samples and noisy data create the first major obstacle to producing a complete set of municipal TFR estimates. In a previous article (Assunção et al. 2005) we proposed a solution to this problem, in the form of EB estimation of local schedules with moving geographic neighbourhoods, and demonstrated its utility with fertility data from Brazil's 1991 Census. Our proposed method also uses EB smoothing, this time for both age-specific fertility and age-specific parity schedules.

Our earlier article and the supplemental web site for this paper (Schmertmann et al. 2012) contain the operational details of the method, so we offer only a brief overview here. Bayesian estimators of all varieties compromise between (1) fitting observed data (e.g., matching the birth/woman ratios in Figure 1), and (2) matching a *prior distribution* that probabilistically describes any likely features of the parameter set that are known before examining the data. In an *Empirical Bayes* procedure, the researcher uses sample data from related domains (e.g., birth/woman ratios from municipalities near Borá) to construct the prior distribution.

Slightly more formally, call the 7×1 vector of true fertility rates for a municipality and suppose that the researcher observes both the local Census sample estimates for those rates (call that vector *Local*) and the vector of average rates in the neighbourhood around the municipality (*Nhood*). Treating all vectors as unknown quantities with probability distributions $P()$, the rules of conditional probability imply that

$$P(\theta | Local, Nhood) = \frac{P(Local | \theta, Nhood) \cdot P(\theta | Nhood) \cdot P(Nhood)}{P(Local, Nhood)} \quad (1)$$

or more simply that

$$P(\theta | Local, Nhood) \propto P(Local | \theta) \cdot P(\theta | Nhood) \quad (2)$$

Equation (2) eliminates terms that do not depend on θ and assumes that the probability of a given local sample, conditional on true local rates θ , is identical regardless of rates in neighbouring municipalities.

The first term on the right-hand side of Equation (2) describes sampling variability: how likely are different *Local* estimates when true rates equal θ ? This probability is lower when θ and *Local* are dissimilar. The second right-hand term describes *a priori* information that the researcher might bring to the problem: how likely are various possible rates θ in neighbourhoods where average rates equal *Nhood*? If rates vary smoothly over space (our key assumption), then this second probability is lower when θ and *Nhood* are dissimilar.

If we further assume that the relevant distributions are approximately multivariate normal, then the logarithms of the two terms in Equation (2) depend on normalised squared-error distances between *Local* and θ , and between θ and *Nhood*. Thus an EB estimator for the vector of age-specific fertility rates in Borá might minimise a scalar function that penalised both deviations from local data and deviations from neighbourhood averages, such as

$$\theta_{EB} = \underset{\theta}{\operatorname{argmin}} \left(\underset{\text{penalised deviation from local estimates}}{(Local - \theta)' \Omega^{-1} (Local - \theta)} + \underset{\text{penalised deviation from neighbours' estimates}}{(Nhood - \theta)' \sum^{-1} (Nhood - \theta)} \right) \quad (3)$$

where Ω matrices and $\sum$ serve as weights representing, respectively, expected sampling noise in the local area and the covariances of age-specific rates across the set of neighbouring municipalities. When the local sample size is large, the matrix Ω is small (i.e., its eigenvalues are small and its determinant is near zero; we will intuitively describe matrices as “big” and “small” in this way). Similarly, $\sum$ is small when the neighbourhood is homogeneous and rates vary little across municipalities.

The solution to the minimisation problem is a matrix-weighted average of local and neighbourhood data

$$\theta_{EB} = Local + \Omega \left[\sum + \Omega \right]^{-1} (Nhood - Local) \quad (4)$$

Note the “shrinkage” property by which estimates for a local area are partially pulled towards estimates from neighbours. Because this estimator operates on the entire vector of local rates, it also borrows strength across age groups within an area (see Assunção et al. 2005 for details).

Our method constructs a unique neighbourhood for each small area. This creates customised subnational areas that are small enough to capture fine geographic variation in demographic rates, but also have sufficiently large samples to produce reliable targets for shrinkage. For each municipality M in the Brazilian data, for example, we selected $\{M + \text{its seven closest municipalities}\}$ as the neighbourhood for M , *unless* this neighbourhood had fewer than 21,000 women age 15–49 in its Census sample. If the $\{M + 7\}$ neighbourhood had fewer than 21,000 women, we added the 8th-closest municipality, 9th-closest, and so on until the extended neighbourhood has at least 21,000 women.

Each small area is the focal point of its own distinct neighbourhood. Neighbourhoods overlap considerably for close-together areas, but are completely distinct when the areas are far apart. In all cases, EB estimators compromise between purely local sample data and the more abundant information from the neighbourhood, with greater use of neighbours' data when the local sample in the focal area is small and when the neighbourhood is homogeneous.

Assunção et al. (2005) contains many technical details on estimating Ω and $\sum$ matrices, on choosing neighbourhood sizes, on interpreting the EB procedure's smoothing properties, and on the performance of the EB estimator with Brazilian census data. The underlying logic of the EB approach, however, is simple: the method averages local and neighbourhood data, with a heavier weight on local data when the local sample is larger.

Figure 2 shows examples of EB smoothing from three Brazilian municipalities with very different sample sizes. Borá appears in Panels A (fertility) and B (parity). Panels C and D contain data for Cajapió, a municipality in the state of Maranhão with 470 women in its

Census sample, very close to the median municipal sample size of $n=467$. Panels E and F present data for Belo Horizonte, one of Brazil's largest cities, with a sample of almost 70,000 women.

Comparison across the three municipalities in Figure 2 illustrates how EB estimators borrow information from neighbours selectively. When the researcher has almost no local information on which to base estimates (i.e., local sample size is small and σ is big), the method borrows heavily from neighbours' data and the EB estimates are close to the neighbourhood average. When the local sample size is very large, as in Belo Horizonte, the EB estimate essentially ignores information from neighbours in favour of the reliable local information. In intermediate cases, such as Cajapió in Panels C and D, the estimated EB schedule often falls between the local and neighbourhood estimates (in the sense explained by Assunção et al. 2005:549–550). Note that the EB estimator often alters the age pattern of the local schedule, as well as its level.

Figure 3 summarises results of EB estimation for the complete set of 5506 municipalities. The horizontal axis in both panels corresponds to direct estimates that use only local birth/woman ratios; the vertical axis represents EB estimates that combine local and neighbourhood data. In the left-hand panel one can see how the EB method reins in TFR outliers, especially those with very low local estimates. The range of EB-estimated TFRs was 1.2–5.9, compared to 0.2–9.2 when using purely local data. The main change caused by EB smoothing of fertility rates is a regression to the mean effect, with low TFR levels pulled upwards toward the (unweighted) municipal mean of 2.4, and high levels pushed downwards. The mean absolute difference between a municipality's EB and Census estimates of TFR was 0.4. Adjustments were largest for locations with the smallest local samples, as suggested by the pattern in the three highlighted municipalities.

The right-hand panel of Figure 3 displays Census and EB estimates for parity among 45–49 year olds. The smoothing method is identical, but EB-Census differences are smaller for parity than for fertility. This occurs mainly because the local Census estimators of cumulative quantities are less noisy and therefore require less smoothing. Although the change induced by EB smoothing is smaller for cohort measures, it is important to note that the method requires stronger prior assumptions about spatial patterns than those necessary for period data. For the cumulative experiences of cohorts in neighbouring areas to be similar, we must assume that neighbouring areas are likely to have similar recent histories, as well as similar present circumstances. We must also assume that migration between small areas is unlikely to cause large changes in local parity distributions.

As with fertility, the EB method alters Census estimates more in small places such as Borá, less in bigger places such as Cajapió, and hardly at all in large places such as Belo Horizonte. Large adjustments to P_7 census values occur only in small municipalities, like Borá, for which average parity at ages 45–49 is inconsistent with parities in other age groups and in nearby places.

Empirical Bayes smoothing is nonparametric. It reduces local sampling variability in the f and P schedules by combining local and neighbourhood sample data, in proportions that vary with local conditions. Another, more traditional way to reduce sampling noise is to assume in advance that all schedules have a particular mathematical form, and then to estimate parameters from local sample data. The relational Gompertz approach to indirect estimation (Booth 1984, Brass 1996, Moultrie et al. 2012) uses the parametric strategy, for example, by assuming that a single pair of parameters (α, β) describes the shape of an area's period and cohort fertility schedules. A parametric model is, in a sense, a very strong *a priori* assumption that one uses in the absence of other data. In the estimation context that we deal

with here such strong assumptions about local rates are unnecessary, because there is a wealth of other (neighbourhood) data from which to borrow strength more flexibly.

Stage 2: A Modified Brass P/F Method

Neighbourhood-based Empirical Bayes smoothing effectively gets the researcher over the first hurdle in municipal-level estimation. It greatly reduces problems caused by small samples and noisy local data, and it increases the signal-to-noise ratio in local fertility rate estimates.

However, EB estimates may still suffer from standard problems of misreporting, particularly various forms of omission that may lead to underreporting of current fertility in a survey or census (United Nations 1983:31–32). Producing an accurate set of several thousand municipal-level estimates requires an automated, reproducible algorithm that will also address potential reporting errors in the EB estimates.

A Modified P/F Method for Changing Fertility Levels

The Brass P/F method is the standard indirect technique for correcting possible errors in the reporting of current fertility (United Nations 1983: Chapter 2). The method uses two key assumptions. First, because of common reporting problems, the Brass approach supposes that parity data (children ever born) are better than fertility data (children born in a recent reference period) for estimating the overall level of fertility. If reported fertility is low (or high) relative to parity, then parity trumps fertility and TFR is adjusted upwards (or downwards). Second, the Brass approach supposes that age-specific fertility rates have been constant over the reproductive lifetimes of women contributing data to the sample. If rates are nearly unchanging, then there should be little difference in a cross-sectional survey between age-specific parities (P_x) and true cumulative period fertility (which we will call φ_x). However, when fertility levels have been changing rapidly, as they clearly were in Brazil prior to 2000, then the relationship between φ_x and P_x becomes more complex. Changing rates make parity-based consistency checks and corrections more difficult to derive (Brass 1996, Feeney 1996, Moultrie and Dorrington 2008).

Our proposed method retains the first Brass assumption, but discards the second. Standard P/F methods are not well suited to circumstances in which fertility levels were changing rapidly over the reproductive lifetimes of women in the sample, which is clearly the case in our Brazilian example (Potter et al. 2002). It is possible to generalise the standard model to allow time-varying rates, however, and to develop a regression-based variant of the Brass P/F approach for adjusting small-area EB estimates. Simulations, reported in the Appendix, suggest that this regression method produces good TFR estimates and is robust to changing fertility.

Start with notation. Suppose that a survey is taken at time $t=0$, and let $\varphi(a)$ denote the true fertility rate at exact age a on that date. Without loss of generality, define rates for other dates via multipliers $K(a,t) \geq 0$, such that

$$\varphi(a,t) = K(a,t) \varphi(a) \quad (5)$$

Call the women who are age x on the survey date *cohort* x ; these women were born at time $-x$, and were a years old at time $a-x$. For the women in cohort x , define

$$\Phi_x = \int_0^x \varphi(a) da = \text{cumulative period fertility (true)} \quad (6)$$

$$F_x = \int_0^x f(a) da = \text{cumulative period fertility (reported)} \quad (7)$$

$$P_x = \int_0^x \varphi(a, a-x) da = \text{parity} \quad (8)$$

$$\mu_x = \frac{1}{\Phi_x} \int_0^x a \varphi(a) da = \text{period mean age of previous childbearing} \quad (9)$$

Denote total fertility at time t as $TFR(t)$, and follow the convention of omitting the time argument when referring to the survey date $t=0$. The demographer's goal is to estimate total fertility on the survey date

$$TFR = \Phi_\omega = \int_0^\omega \varphi(a) da \quad (10)$$

from a set of reported rates $f(a)$. The Brass method uses survey data to adjust cumulative reported fertility F :

$$Y_x = F_\omega \cdot \frac{P_x}{F_x} \quad (11)$$

which generally yields different estimates when using P and F data from different cohorts x .

If reporting errors are nearly equiproportional across ages [i.e., if there is a single constant c such that $f(a) \approx c \cdot \varphi(a)$], Equation (11) becomes

$$Y_x \approx \Phi_\omega \cdot \frac{P_x}{\Phi_x} = TFR \cdot \frac{P_x}{\Phi_x} \quad (12)$$

In terms of past and present rates, this expression is

$$Y_x \approx TFR \cdot \frac{\int_0^x K(a, a-x) \cdot \varphi(a) da}{\int_0^x \varphi(a) da} = TFR \cdot \bar{K}_x \quad (13)$$

where K_x is the average ratio of past to present fertility, weighted by *period* rates up to age x . Approximating the mean of the age-specific function $K(a, a-x)$ by the function's value at the mean age $a = \mu_x$ simplifies the expression further:

$$Y_x \approx TFR \cdot \bar{K}_x \approx TFR \cdot K(\mu_x, \mu_x - x) \quad (14)$$

At this point one needs a model of fertility change – i.e., a functional form for $K(a, t)$. The strongest version of the Brass method assumes unchanging fertility rates, so that $K(a, t) = 1$ at all ages and times. Under that strong assumption, adjustments Y_x using the P/F ratio at any age x will produce the correct value for period TFR. In practice, demographers usually use a weaker version, hoping that fertility has been approximately constant over the reproductive lifetimes of the *young* women in the survey. In this case one calculates Y_x only for young age groups, such as 20–24 or possibly 25–29 year olds. Such an approach implicitly assumes that $K(a, t) \approx 1$ for young ages and recent times, so that quantities like $K_{20-24} \approx 1$.

A regression approach using P/F adjustment

We propose a simple generalisation of the unchanging rate assumption. In particular, suppose that, over the reproductive lifetimes of the women surveyed, fertility rates have been changing at a constant rate that is identical for all ages

$$K(a, t) = e^{rt} \quad (15)$$

Under this assumption total fertility also changes exponentially:

$$TFR(t) = e^{rt} \cdot TFR \quad (16)$$

Note that the strong Brass model is a special case in which $r=0$.

In the generalised model the P/F adjustments in Equation (14) become

$$Y_x \approx TFR \cdot K(\mu_x, \mu_x - x) \approx TFR \cdot e^{r(\mu_x - x)} \quad (17)$$

or alternatively

$$Y_x \approx TFR(\mu_x - x) \quad (18)$$

These equations have important implications. Equation (18) shows that P/F estimates from different ages or age groups may be considered as *good estimates of past fertility*, rather than as *bad estimates of current fertility* [cf. Equation (13)]. P/F-adjusted total fertility for women age x approximately equals period TFR at the average time of previous childbearing in that group.

Feeney (1996) presented a reinterpretation of P/F adjustment that followed a similar logic. However, he proposed time allocation based on the difference between current age and unconditional (i.e. past and future) mean time of cohort childbearing, $\mu - x$, rather than on the conditional (past only) mean $\mu_x - x$. Equation (17) makes it clear that the correct time allocation uses μ_x . Moultrie and Dorrington (2008) tested Feeney's method in simulated histories using Feeney's $\mu - x$ rather than $\mu_x - x$ as the basis for time allocation; the poor results that they report for Feeney's method may result from this difference.

Equation (17) suggests a regression approach that uses time-allocated values to estimate both current TFR and the recent rate of fertility change. Specifically, because

$$\ln Y_x = \ln \left[F_\omega \frac{P_x}{F_x} \right] = \ln TFR + r \cdot (\mu_x - x) \quad (19)$$

linear regression of the logarithms of P/F-corrected TFRs on $\mu_x - x$ produces an easily interpretable intercept (the logarithm of current TFR) and slope (the rate of fertility change).

When data are in five-year age groups, the researcher faces two practical decisions in implementing the regression model: (1) how to calculate age-group averages for cumulative fertility F and for elapsed times since previous births ($\mu_x - x$); (2) how to weight the least-squares regression.

We propose a penalised spline interpolation procedure, described in detail in a working paper (Schmertmann 2012: Appendix), for resolving the first issues. This method uses a set of predetermined constants to construct estimated $(F_{15-19} \dots F_{45-49})$ and $(\mu_{17.5-17.5} \dots \mu_{47.5-47.5})$ values from $(f_{15-19} \dots f_{45-49})$.

The choice of regression weights is important because estimated P/F ratios on the left-hand side of Equation (19) have high variance for young ages with low F. One can approximate these variances logically, as follows. If cumulative fertility through age x equals F_x , then $B_x \sim \text{Poisson}(n F_x)$ is a reasonable model for the number of children ever born to a group of n x -year-old women. In that case the mean parity in a sample, $F_x = n^{-1} B_x$, would have mean $E(F_x) = F_x$ and variance $V(F_x) = n^{-1} F_x$. Similarly, if period births before and after age x are statistically independent then $\text{cov}(F_x, F) = n^{-1} F_x$. Combining these derivations with standard delta-rule approximations $V(\ln y) \approx \sigma_y^2 / \mu_y^2$ and $\text{cov}(\ln x, \ln y) \approx \sigma_{xy} / (\mu_x \mu_y)$ yields variances for the terms on the left-hand side of regression equation (19):

$$V(\ln Y_x) \propto P_x^{-1} + F_x^{-1} - F_\omega^{-1} \quad (20)$$

When fitting the model in Equation (19), one can use the inverses of these approximate variances as regression weights. Following precedent (e.g., United Nations 1983:35) we suggest omitting 15–19 year olds entirely; this is the same as using a zero regression weight for the first age group. If parity reports from the women 40+ or 45+ are also considered unreliable, they can be omitted or given zero weight, but we retained them for our analysis of the Brazilian data.

The assumptions leading to Equations (18) and (19) are quite strong, but testing (described in the Appendix) demonstrates that a regression approach works very well. In simulated fertility transitions that do not conform to the simple assumption in equation (15), a weighted, regression-based P/F model nonetheless produces TFR estimates that have smaller average errors than the customary P_2/F_2 or P_3/F_3 estimators.

Regression Examples

Table 2 contains the regression data for our example mid-sized municipality, Cajapió. EB-smoothed values for f_i and P_i serve as inputs to the spline interpolation for cumulative fertility and time-since-birth averages F_i and μ_i . These moments, in turn, allow calculation of P/F-adjusted total fertilities $Y_i = F(P_i/F_i)$ and allocated times $\mu_i - (12.5 + 5i)$, as well as weights from Equation (20).

Cajapió's data exhibit P/F ratios well above unity in all age groups. This suggests significant underreporting of current fertility in the census. In addition, with the exception of the youngest women, there is a monotonic increase in P/F ratios as one moves to older age groups. The pattern of increasing ratios is evidence of rapidly decreasing fertility in this municipality: P/F-adjusted TFRs range from $Y_2 = 3.93$ using P_2/F_2 (which corresponds to TFR approximately 2.9 years before the census) to $Y_7 = 7.31$ (21.4 years before the census).

The regression model uses adjusted TFR levels from $i=2 \dots 7$, together with their time allocations and inverse-variance weights, to arrive at an estimated trend $\ln TFR(t) = 1.207 - 0.037t$. (Note that the usual Brass approach, $TFR(0) = Y_2$, is a special case of this regression, in which $t=0$ by assumption and only the second age group has positive weight.) The modified Brass estimate of total fertility in Cajapió at time $t=0$ (i.e., approximately six months before the census date) is therefore $e^{1.207} = 3.34$. This represents a significant upward correction from the EB estimate of 2.69, and suggests that reference period or other reporting errors are an important source of downward bias in Census data for Cajapió.

Note that the standard Brass adjustment for this bias, using $Y_2 = F (P_2/F_2) = 3.93$, would overestimate fertility in the reference period. This occurs primarily because parity among 20–24 year olds in the second age group includes higher fertility rates from the (recent) past that tend to make $P_2/F_2 > 1$ even in the absence of reporting errors. Our approximation demonstrates that this effect is non-negligible when, as in contemporary Brazil, fertility is decreasing rapidly and childbearing at young ages is relatively high. In the case of Cajapió, one would expect $Y_2 = F (P_2/F_2)$ to equal period TFR approximately 2.9 years earlier; thus at 3.7% annual decline Y_2 will be approximately $e^{-0.037(-2.9)} = 1.113$ times the current TFR. In other words, even with perfect data the usual Brass P_2/F_2 approach would overestimate current TFR by around 11% in this situation, because of rapidly falling fertility.

The calculations in Table 2 also highlight statistical concerns with the usual P_2/F_2 approach. The inverse-variance regression weights illustrate that Y_2 estimates are far less precise than estimates from older age groups; this occurs because it is very difficult to estimate the inverse of a small number, like $1/F_2$, accurately. It is therefore unclear whether the usual approach of ignoring the (more temporally biased but much less variable) Y values from older age groups in favour of (less biased but much more variable) Y_2 and Y_3 will yield a superior estimator of current TFR.

A final reason for concern with the usual P_2/F_2 calculation is that underreporting of births to teenage mothers (f_1) causes more upward bias in $1/F_2$ than in $1/F$ ratios for older age groups. Thus the Brass estimator $TFR = Y_2$, even if allocated properly in time, might suffer more positive bias from teen under-reporting than alternative estimators, like our regression approach, that use information from a mix of age groups.

Figure 4 illustrates the Brass regression procedure for our three example municipalities with small, medium, and large sample sizes. In each panel, six points represent the time-allocated TFR estimates $exp(Y_i)$ for age groups $i=2\dots7$, and the solid line represents the fitted regression estimate for the time path of total fertility. The data in Table 2 were used for the Cajapió regression in the middle panel. The C and E labels on the right-hand vertical axes represent the local Census (C) and Empirical Bayes (E) estimates of municipal TFR in the year prior to the census (labelled as “2000”). The height of the fitted regression line in year 2000 is the Brass regression estimate of current TFR. Points in each panel correspond to time-allocated TFR estimates for age groups $i=2\dots7$, labelled with the youngest age in each five-year group.

As with the EB procedure in Figure 2, the regression procedure generally smoothes more in small-sample cases. All three examples in Figure 4 exhibit P/F patterns that are strongly consistent with falling fertility, and which conform well to the model of exponential change over the reproductive lives of surveyed women. Fertility decline over 1975–2000 in Cajapió, in North-eastern Brazil, appears to have been much more rapid (–3.7% per year) than in the South-eastern municipalities of Borá and Belo Horizonte (–1.5% and –2.3%, respectively). These differences in the rate of decrease are consistent with earlier studies of Brazilian fertility decline at slightly larger subnational scales (Potter et al. 2010).

Regression-based P/F adjustments to TFR tend to be positive, suggesting varying levels of underreporting: Borá’s EB estimate of 1.94 rises to 2.21, for example, while Cajapió’s EB estimate of 2.69 changes to 3.34. Belo Horizonte’s TFR estimate falls very slightly with P/F correction, from 1.68 to 1.63, which suggests very slight overreporting of current fertility.

Generally speaking, the P/F regressions and TFR adjustments illustrated in Figure 4 are representative of the patterns, fits, and levels found in Brazil’s 2000 Census data. We have thus far presented our two-stage methodology (EB smoothing of local data, followed by P/F

regression adjustment) via detailed examples. We turn in the next section to a broader analysis of the complete set of 5506 municipal-level calculations.

EBB Results

We begin with a comparison of the map of local Census (i.e., weighted births/woman) estimates to the map of EBB estimates. Figure 5 contains Census estimates in the left-hand map, and Empirical Bayes plus Brass (EBB) estimates on the right. Both maps have a general south-to-north gradient of increasing fertility, with additional areas of high fertility in the inland areas of North-eastern Brazil. Both maps also indicate that fertility is below replacement level in many areas of Brazil's Southeast and in the state of Goiás.

The EBB map, however, is less mottled and smoother-looking than the Census map. This increase in smoothness arises mainly because the EBB estimator purges some of the sampling noise that makes many Census TFR estimates unreliable. In areas with small populations and small samples, EBB borrows information from presumably similar surrounding municipalities, in order to produce estimates with much lower mean absolute errors. The result is a smoother map that is more likely to represent subnational fertility patterns accurately.

Summary statistics in the maps' subtitles illustrate these differences numerically. Across pairs of adjacent municipalities, the mean absolute difference in neighbours' EBB estimates (MADN) is .38 children/woman; the same quantity for the Census map is .61. Moran's I , a measure of correlation between neighbouring values, is nearly identical for the two methods, with a value of .73.

A good small-area method should not generate many implausible outliers. Figure 6 addresses this point by showing the joint distributions of Census and EBB estimates across municipalities (top panels), and the marginal distributions of each set of estimates (bottom panels). Municipalities are disaggregated by the size of their 2000 Census samples: the 500 municipalities with the smallest local samples (sample sizes n in these areas are between 34 and 150 women) appear in the left-hand panels, the 500 with the largest samples in the right-hand panels (n from 1449 to 314732), and all other municipalities in the central panels.

Figure 6 illustrates several desirable features of the EBB approach. First, the EBB estimator virtually eliminates all implausibly low TFR estimates. For instance, many small and medium-sized municipalities have Census TFR estimates below 1.30. The EBB procedure treats the small-area census results almost entirely as sampling error, and (in order to better conform with municipal parity and neighbourhood fertility data) raises local TFR estimates to much more plausible levels. This effect is most apparent in the regression-to-the-mean pattern in the two leftmost panels of Figure 6: for the smallest areas, EBB treats very low Census TFR estimates almost entirely as sampling noise, raising all these areas' TFR estimates to at least 1.5.

A second desirable property evident in Figure 6 is the EBB estimator's differential treatment of areas with different-sized samples. Although the EBB algorithm is identical for each municipality, the marginal distributions of the estimators in the three bottom panels illustrate that the effect of EBB depends largely on sample size. Recall that EBB is a two-step procedure: Empirical Bayes shrinkage, followed by (usually upward) Brass-style parity adjustment.

In the smallest municipalities, the shrinkage effect of the Empirical Bayes step appears to dominate: variability falls, and there is only a small increase in the mean and modal TFR estimates. The net effect in these smallest areas is an EBB TFR distribution that is

considerably more concentrated around the mode, with far fewer low and high values, than the census estimates.

In medium-sized municipalities, in contrast, the net effect of shrinkage and parity adjustment is more complex: EBB tends to increase very low census estimates (both shrinkage and parity adjustment tend to raise low estimates), but tends to change high estimates less (for these estimates the downward shrinkage effects and upward parity effects are roughly in balance. The result is a more right-skewed EBB distribution, with a slightly higher mean and mode than the census TFRs.

In the largest municipalities, the effects of EBB are qualitatively similar to those in the medium-sized areas, but they differ greatly in magnitude. In large municipalities EB shrinkage effects are very small, because local samples are large and local estimates are therefore far less noisy. Parity adjustment effects for large areas are, on average, very small and positive. The result is a distribution of large-area EBB estimates that is nearly identical to the distribution of census estimates.

Comparison of EBB to Alternative Estimators

It is instructive to compare the EBB estimator with alternative sets of municipal-level estimates. Table 3 describes several possibilities, and shows their national population-weighted means. The first row, labelled *Census*, represents direct estimators that calculate local births/woman by age group and sum those rates to calculate local total fertility. This method does not address the high variability of estimates for small municipalities, nor does it use parity data as a consistency check. Census estimators are sensitive to sampling error, and prone to generate erroneously low TFR values for many municipalities. The 95% intervals in the penultimate column of Table 3 illustrate these problems.

Empirical Bayes spatial smoothing, on the second row of Table 3, narrows the distribution of TFR estimates, pulling both low and high extremes toward the mean. This is the regression-to-the-mean effect seen earlier in Figure 3. EB smoothing improves the overall set of estimators, in particular by eliminating implausibly low TFR estimates in small municipalities. However, without parity corrections for possible misreporting it is likely that in many locations the EB estimates are too low.

EBB estimates, which appear on the table's third row, add parity correction to complete our two-step procedure. The distribution of estimates shifts in different ways depending on municipal sample sizes, as illustrated in Figure 6. Overall, Bayes regression correction raises the national average slightly from the Census and EB levels – population-weighted TFR rises from near 2.2 to 2.3.

The last row of Table 3 summarises TFR estimates from the United Nations Development Programme's *Human Development Atlas* (UNDP 2003). These estimates were generated by an alternative procedure that uses two distinct estimation algorithms (details in Portuguese, Horta et al. 2005). For municipalities with at least 30,000 residents, the UNDP estimates are Brass P_2/F_2 -corrected TFR calculations that use only local data. For smaller municipalities, the UNDP method estimated local TFR in two steps: (1) calculate Brass P_2/F_2 estimates of TFR at a higher level of Census geography, for collections of municipalities known as *microregions*, and (2) multiply the microregion's TFR by the ratio of local (P_2+P_3) over microregional (P_2+P_3).

This UNDP strategy for reducing small-area sampling variance is sensible, but ad hoc. TFR estimates will differ for two municipalities with identical local data – if one has 29,999 residents and the other has 30,000. Furthermore, using a rigid hierarchical Census geography

(rather than spatial neighbourhoods) may create spatial seams in the estimates, because pairs of adjacent small areas with similar local data could have very different TFR estimates if they belong to different microregions.

Despite potential difficulties with ad hoc adjustment of small-area data, the main concern with the UNDP approach is its reliance on P_2/F_2 . As we have demonstrated earlier in this paper, this standard Brass correction will tend to overestimate fertility levels by perhaps 5–10% when fertility is falling rapidly. Figure 7 provides a final summary of the various municipal estimators in Table 3, aggregated up to the level of Brazil's twenty-seven states for easier visual comparison. At the state level UNDP estimates are consistently higher than EBB estimates, by approximately .15 to .20 children per woman, and are probably overestimates due to the temporal bias in the P_2/F_2 correction.

Discussion and Conclusion

In this paper we have combined old and new methods to tackle a thorny problem. Producing estimates for a very large set of places from sparse data requires automated, reproducible algorithms that address some familiar challenges – namely, high sampling variability and possible errors by survey respondents. Our EBB approach attacks these two problems sequentially, first smoothing survey data by borrowing strength across age groups and spatial neighbourhoods, and then applying an appropriate variant of parity correction to the smoothed local fertility rates.

Assessing a set of more than 5000 estimates is difficult, particularly when the true values of the estimands are unknown and unavailable from other sources. Given these difficulties, empirical evaluation can only be undertaken in broad terms, as in the previous subsection, to see if a method produces sensible results and avoids obvious errors. In this sense, EBB results appear to be quite good.

In addition to empirical testing, one can also consider general criteria that characterise good estimation methods. We argue that the EBB method is a good approach to estimating small-area rates in countries with limited vital registration because it:

- is 100% automated and reproducible
- uses an identical estimation algorithm for every area
- reduces the effect of sampling variance in areas with small populations (e.g. Borá)
- yields results similar to current best-practice estimators in areas with large populations (e.g. Belo Horizonte)
- uses parity data to adjust for possible reporting errors in current fertility
- uses parity data in a way that is consistent with changing fertility levels

Our EBB method requires census data identifiable by age and by the geographic or administrative unit of interest. It is most useful when vital registration is incomplete, and when small local fertility samples make it difficult to estimate rates reliably. While Brazil was one of the first countries to meet both criteria, it is rapidly approaching complete birth registration. Complete birth registration is, however, still a relatively rare phenomenon among countries in Africa, Latin America, and many parts of Asia.

The regression variant of parity correction works best when fertility change has been steady and in one direction over a period of several decades before the survey. Even when previous fertility change has not followed that assumed pattern, the Appendix simulation demonstrates that the variance reduction from using a P/F regression with multiple data

points, rather than only one or two as in the customary Brass approach, may more than compensate for the resulting bias.

The EBB estimation approach produces estimates at a finer level of geographic detail than would otherwise be possible from census or survey data. This information can be important: it is clear from Figure 5, for example, that state-level aggregation would obscure many features of Brazil's current fertility patterns. We believe that two major groups of demographic consumers can benefit from additional detail provided by EBB estimation. Academic demographers gain richer perspective on the processes underlying fertility change from more geographically detailed data (Galloway et al. 1994, Brown and Guinnane 2002, Potter et al. 2002, Potter et al. 2010). Public health professionals and policy makers gain valuable information that allows more precise targeting of program expenditures, and improved local population forecasts.

Standardising the EBB method for use by governments and local agencies requires further work. We have begun the process by posting complete data and *R* code on our project web site. However, it would be valuable to further develop software to improve ease of use.

When vital registration is incomplete, demographic surveys and census data must fill the breach. In most countries even very large national samples will not produce sample sizes big enough to produce good estimates at fine levels of geography. In the long run, vital registration systems may improve enough to obviate the need for model-based methods. In the short run, however, an increasing demand for small-area estimates and projections will likely lead 21st-Century demographers to rediscover some of their 20th-Century roots and, as we have done in this paper, to develop new methodological variations on old themes.

Acknowledgments

This research was supported by grant R01 HD41528 from the US National Institutes of Health, Institute of Child Health and Human Development. We thank the editor and two anonymous reviewers for insights and suggestions that improved the article significantly. We also thank participants at the June 2011 IUSSP seminar on "Current Issues and Frontiers in Demographic Research", held in honour of Prof. José Alberto Magno de Carvalho, for helpful comments on an earlier draft.

References

- Assunção RM, Schmertmann CP, Potter JE, Cavenaghi SM. Empirical Bayes estimation of demographic schedules for small areas. *Demography*. 2005; 42(3):537–558. [PubMed: 16235612]
- Booth H. Transforming Gompertz's function for fertility analysis: the development of a standard for the relational Gompertz function. *Population Studies*. 1984; 38(3):495–506. [PubMed: 22087671]
- Brass W. Demographic data analysis in less developed countries: 1946–1996. *Population Studies*. 1996; 50(3):451–467.
- Brass, W.; Coale, AJ.; Demeny, P.; Heisel, DF.; Lorimer, F.; Romaniuk, A.; Van De Walle, E. *The Demography of Tropical Africa*. Princeton University Press; 1968.
- Brown JC, Guinnane TWG. Fertility transition in a rural, Catholic population: Bavaria 1880–1910. *Population Studies*. 2002; 56(1):35–50. [PubMed: 12102098]
- de Carvalho JAM. Regional trends in fertility and mortality in Brazil. *Population Studies*. 1974; 28(3): 401–421. [PubMed: 22077734]
- Cliff, AD.; Ord, JK. *Pion*. 1981. *Spatial Processes*.
- Coale AJ, Trussell TJ. Model fertility schedules: variations in the age structure of childbearing in human populations. *Population Index*. 1974; 40(2):185–258. [PubMed: 12333625]
- Feeney, G. [Accessed 30 May 2010] A new interpretation of Brass's P/F ratio method applicable when fertility is declining. 1996. Research note posted on <http://tinyurl.com/feeney-pf>

- Galloway PR, Hammel EA, Lee RD. Fertility decline in Prussia, 1875–1910: a pooled cross-section time series analysis. *Population Studies*. 1994; 48(1):135–158. [PubMed: 11639272]
- Horta CJG, de Carvalho JAM, Nogueira OJO. Evolução do comportamento reprodutivo da mulher brasileira 1991–2000: cálculo da taxa de fecundidade total em nível municipal [Evolution of the reproductive behaviour of Brazilian women 1991–2000: calculation of total fertility at the municipal level]. *Revista Brasileira de Estudos Populacionais*. 2005; 22(1):131–140.
- Instituto Brasileiro de Geografia e Estatística (IBGE). National Demographic Census. Rio de Janeiro: IBGE; 2000.
- Instituto Brasileiro de Geografia e Estatística (IBGE). Estatísticas do Registro Civil, comentário. 2008; 35 <http://tinyurl.com/ibge-subregistro>.
- Moultrie TA, Dorrington RE. Sources of error and bias in methods of fertility estimation contingent on the P/F ratio in a time of declining fertility and rising mortality. *Demographic Research*. 2008; 19(46):1635–1662.
- Moultrie, TA.; Dorrington, RE.; Hill, AG.; Hill, KH.; Timæus, IM.; Zaba, B., editors. Tools for Demographic Estimation. Paris: International Union for the Scientific Study of Population; 2012. <http://demographicestimation.iussp.org/> [last accessed 17 July 2012]
- Potter JE, Schmertmann CP, Assunção RM, Cavenaghi SM. Mapping the timing, pace, and scale of the fertility transition in Brazil. *Population and Development Review*. 2010; 36(2):283–307. [PubMed: 20734553]
- Potter JE, Schmertmann CP, Cavenaghi SM. Fertility and development: evidence from Brazil. *Demography*. 2002; 39(4):739–761. [PubMed: 12471852]
- Schmertmann CP. A system of model fertility schedules with graphically intuitive parameters. *Demographic Research*. 2003; 9(5):81–110.
- Schmertmann, CP. MPIDR Working Paper WP-2012–022. Rostock: Max Planck Institute for Demographic Research; 2012. Calibrated spline estimation of detailed fertility schedules from abridged data. <http://tinyurl.com/calibrated-spline>
- Schmertmann, CP.; Cavenaghi, SM.; Assunção, RM.; Potter, JE. 2012. Supplemental website for this paper at <http://schmert.net/BayesBrass>
- United Nations. Manual IV. Methods of estimating basic demographic measures from incomplete data. United Nations Publications; 1967. Sales No. 67.XIII.2
- United Nations. Manual X: Indirect Techniques for Demographic Estimation. United Nations Publication; 1983. Sales No. E.83.XIII.2
- United Nations Development Programme; Atlas of Human Development in Brazil, editor. [last download 26 Sep 2011] Atlas do Desenvolvimento Humano no Brasil. 2003. Software at <http://www.pnud.org.br/atlas>

Appendix: Comparing Estimators over a Simulated Fertility Transition

Likely Patterns of Bias

It is important to understand whether alternative methods of estimating period TFR are robust to violations of their principle assumptions. The standard P_2/F_2 correction assumes that fertility rates at ages below 25 have been constant over the reproductive lifetimes of women 20–24. This assumption is best satisfied before the start of transition (when all rates are constant), and for the period starting about 10 years after transition's end. At other times, when transition is underway or has only very recently ended, we would expect a positive bias in the P_2/F_2 correction. As discussed in the text, P_2/F_2 bias is likely on the order of +5 to 10% during a typical transition.

Our regression method, in contrast, assumes that fertility rates have been falling at a constant exponential rate, identical at all ages, over the reproductive lifetimes of women 15–49. Because fertility is nearly complete in most populations by age 40, this effectively means that the regression approach assumes that *fertility has been falling steadily over the last 25 years before the sample date*. This assumption will be satisfied fairly well in the

middle and late stages of a long transition. However, early in a transition it will not hold, because the parities of older women are unaffected by secular decline. Early in a transition P_i/F_i in older groups will thus be lower than anticipated by a constant-decline model, and the regression method will *underestimate* the rate of decline and *overestimate* the current TFR. (To see this, imagine the effect of a downward shift in the “40” and “45” points on the regressions in Figure 4.) After a transition ends, the opposite holds true: P_i/F_i in older groups will be higher than anticipated, the regression method will overestimate the rate of decline, and it will *underestimate* current TFR. Our regression method is therefore likely to overestimate TFR during early stages of a transition, and to underestimate for several decades after a transition has ended.

Transition Model and Monte Carlo Sampling

Once one knows the expected strengths and weaknesses of the methods in principle, the next step is to measure them in a realistic example. In order to understand the costs and benefits of the regression method more fully, we compared standard and regression-based P/F correction during a simulated fertility transition that did not match the assumptions of either model. In this experiment we defined an archetypal transition, in which TFR is constant at 6.0 for many years, falls to 2.5 over a 40-year period, and then remains constant at 2.5. We chose this transition length, these TFR levels, and this age pattern of change to approximate the Brazilian situation. However, experiments with alternatives (not shown) satisfied us that our main conclusions are insensitive to these specific assumptions.

Figure A1 illustrates the changing period ASFR schedules over the course of the simulated transition, which are quadratic spline schedules (Schmertmann 2003) with linear changes in parameters as described in the figure caption. Time runs over integer years $t = -35, -34, \dots, +75$. Before $t=0$, the fertility schedule is unchanging at the upper dark curve in Figure A1, and TFR is 6.0. Over the period $t = 0 \dots 40$ the schedule changes as illustrated, finally reaching the lower dark curve and remaining constant afterwards, at a TFR of 2.5. This sequence of schedules $f(a, t)$ violates the assumptions of our regression model, because the shape of the period schedule is not constant. In the simulation, as in Coale-Trussell (1974) and other models of fertility decline, change is concentrated at older maternal ages, so that both the period mean and period standard deviation of the age at childbearing decrease as the transition progresses.

We suppose that when a sample is taken at any time t_s , it includes N women in each single-year age cohort $x=12 \dots 49$. For each cohort we generated a sequence of random birth counts $B_{at} \sim \text{Poisson}[N f(a, t)]$ over the relevant (a, t) combinations, and then recorded the corresponding parity ${}_1P_x$ and period fertility ${}_1f_x$ schedules. Aggregating into five-year age groups produced a sample of $({}_5P_x, {}_5f_x)$ data, from which we estimated $\text{TFR}(t_s)$ using both the regression and the standard P_2/F_2 method (United Nations 1983:31–32). For the purposes of this exercise we assumed accurate reporting of period fertility, so that any errors arise from changing rates rather than from differential misreporting by age. Repeating the process generated 500 independent TFR estimates for each method in each year from $t_s = -5$ through $+75$.

Simulation Results

Figure A2 shows the quantitative results from the Monte Carlo experiments with our simulated transition. The top two panels (A and B) correspond to finite sample performance with $N=100$ women per single year of age, which is roughly comparable to a DHS or other large national survey. The figure's bottom panels (C and D) show the limiting behaviour of the estimators as sample sizes grow arbitrarily large and sampling variances go to zero. The

left-hand panels A and C illustrate root mean squared errors (which incorporate both bias and sampling variance); the right-hand panels B and D illustrate bias.

Bias results for both small and large samples in right-hand panels B and D show the expected patterns. During the transition, the regression method exhibits lower bias, with a particularly large advantage over the P_2/F_2 method if the transition has been underway for more than 20 years. After a transition ends, P_2/F_2 bias drops from about 7% to zero within about five years. During the post-transition period the regression method's bias changes from near zero to approximately -7% over about ten years, and eventually returns to zero in about 25 years.

Root mean squared error (RMSE) results appear in left-hand panels A and C. RMSE provides no additional information for large samples (panel C), because variance goes to zero and RMSE therefore equals absolute bias. Panel A, showing RMSE for finite samples, is much more important. Because P/F ratios are imprecisely estimated, the P_2/F_2 method (which uses only one P/F ratio as input) is much less stable than the regression method (which uses six). Consequently, for samples of size $N=100$ in this particular simulation, the expected error of the P_2/F_2 estimator is higher than that of the regression estimator at all times – before, during, and after transition. During the late stages of transition, average errors in P_2/F_2 -estimated TFR are between two and three times larger than regression errors.

Based on these simulation results, the regression method outperforms the usual P_2/F_2 correction. Panel A shows that in finite samples that are much larger than one usually has for small geographic areas, typical errors are likely to be smaller for the regression method during all phases of a transition – and even when fertility rates are constant. If rates have been declining for decades, the regression method's biases and average errors are much smaller than those of the P_2/F_2 approach.

Figure A1.

Simulated period fertility schedules during a 40-year transition. Initial TFR=6.0 and final TFR=2.5. Curves are generated from the four-parameter quadratic spline family (Schmertmann 2003), with linear changes in parameters over a 40-year period: maximum fertility level R changes from .30 to .17; initial age changes from 12 to 14; peak age P remains constant at 26; halfway point of right-hand side H changes from age 39 to 34. In this transition, average annual rates of decrease at ages (20, 25, 30, 35, 40) are (-1.9, -1.4, -1.8, -2.7, -3.7)%. *Source:* Authors' calculations.

Figure A2.

Root mean squared error (RMSE) and bias of TFR estimates using regression and P_2/F_2 correction – from samples taken at different stages of a simulated transition. The transition occurs during the shaded period, with TFR falling steadily from 6.0 to 2.5 over the period from $t=0$ to 40 on the horizontal axis. Top panels show errors for samples with $N=100$

women in each single-year age group; bottom panels are for the limiting case as N and sampling variance approaches zero. All errors are measured as a percentage of period TFR. Curves in the $N=100$ cases are smoothed (dots represent results from 500 Monte Carlo samples drawn for each transition year). *Source:* Authors' calculations.

Figure 1. 2000 Census $5f_a$ estimates (births last year/women) for the municipality of Borá (circles), for all of Brazil (dark line), and for a set of 34 municipalities around Borá (grey line). *Source:* IBGE (2000) and authors' calculations.

Figure 2.

Census estimates (open circles), neighbourhood averages (grey lines), and Empirical Bayes estimates (dark line with points), for three example municipalities with different sample sizes. Age-specific fertility in left-hand panels A, C, and E; parities in right-hand panels B, D, and F. *Source:* As in Figure 1.

Figure 3. Census versus Empirical Bayes (EB) estimates for 5506 Brazilian municipalities in 2000. Each point corresponds to a municipality, with three example locations from the text highlighted. Total fertility in left-hand panel; parity at ages 45–49 in right-hand panel. Census and EB estimates are equal along the diagonal lines. *Source:* As in Figure 1.

Figure 4.

Brass P/F regression adjustment for three example municipalities. Vertical positions of C, E, and the regression intercept at year 2000 represent TFR_{Cens} (Census estimate), TFR_{EB} (Empirical Bayes), and the modified Brass estimate for TFR, respectively. Points are labelled by the youngest age in a five-year age group. See text for details. *Source:* As in Figure 1.

Figure 5.

Census and Empirical Bayes plus Brass (EBB) estimates of municipal total fertility, 2000. State boundaries in black. MADN = mean absolute difference between neighbours, averaged over 16419 unique pairs of adjacent municipalities. Moran's I measures correlation between values in adjacent municipalities (Cliff and Ord 1981, p. 17). *Source:* As in Figure 1.

Figure 6.

Census and EBB estimates of 2000 total fertility, by size of municipality. Left panels contain the 500 municipalities with the smallest census sample sizes; right panels contain the 500 largest samples; middle panels contain the remaining 4506 municipalities. Scatter plots in the top panels show pairs of alternative estimates, with one point per municipality. Densities in the lower panels illustrate differences between the marginal distributions of Census and EBB estimates within each size class. Note that horizontal scales in the top and bottom panels differ. *Source:* As in Figure 1.

Figure 7.

Alternative estimation methods for municipal fertility, aggregated to state level using population weights, Brazil 2000. Methods are direct estimates from Census (bars), Empirical Bayes spatial smoothing (dashed line with x), Empirical Bayes plus Brass regression correction (solid line), UN Development Programme (UNDP 2003 – dashed line). Population-weighted rates of TFR change, estimated from regressions and written as percentages, appear at the base of each state's column. States are sorted in descending order of Census TFR estimates. State abbreviations at <http://tinyurl.com/br-states>. *Source:* As in Figure 1.

Table 1

Sample size distributions for women and births last year across 5506 municipalities, Brazil 2000 census microdata. *Source:* IBGE (2000).

Age Group	# Sampled Women			# Sampled Births Last Year		
	5%ile	Median	95%ile	5%ile	Median	95%ile
15–19	23	99	464	0	7	38
20–24	19	80	415	2	11	56
25–29	17	67	358	1	8	40
30–34	17	64	340	0	5	26
35–39	16	59	331	0	2	14
40–44	14	49	279	0	1	5
45–49	12	43	231	0	0	1
15–49	123	467	2365	7	36	174

Table 2

2000 Census data for five-year age groups $j=1 \dots 7$, Cajapió. *Source:* IBGE (2000) and authors' calculations.

<i>i</i>	1	2	3	4	5	6	7
Ages	15–19	20–24	25–29	30–34	35–39	40–44	45–49
<i>Empirical Bayes Estimates</i>							
P_{EB}	.32	1.46	2.67	3.64	4.83	6.30	7.28
f_{EB}	.105	.179	.113	.062	.047	.026	.005
<i>Interpolated Moments</i>							
F_i	.14	1.00	1.75	2.16	2.43	2.61	2.68
μ_i	16.3	19.6	21.8	23.3	24.6	25.7	26.1
<i>Regression Inputs</i>							
(H/H_i)	2.21	1.46	1.53	1.69	1.99	2.41	2.72
$(Y_i - F_i) / (H/H_i)$	5.93	3.93	4.10	4.53	5.34	6.47	7.31
Avg $(\mu_k - x)_{ij}$	-1.2	-2.9	-5.7	-9.2	-12.9	-16.8	-21.4
Weight	0	.76	1.74	2.74	4.05	5.90	7.20
<i>EB TFR</i>							
F = 2.69							
<i>Standard Brass TFR</i>							
$Y_2 = F \quad (P_2/F_2) = 3.93$							
<i>Brass Regression</i>							
Ln $Y_i \approx 1.207 - 0.037 \text{ [Avg}(\mu_k - x)_i]$							
EBB TFR = $e^{1.207} = \mathbf{3.34}$							

Table 3

Distribution of municipal TFR estimates for alternative methods across 5506 Brazilian municipalities, 2000.
Source: IBGE (2000), authors' calculations, UNDP (2003).

Estimator	Smoothing	Parity Adjustment	2.5% to 97.5% ile	Weighted Mean ¹
Census	None	None	1.16 – 4.38	2.21
Empirical Bayes (EB)	Spatial Smoothing	None	1.66 – 3.86	2.19
Empirical Bayes + Brass (EBB)	Spatial Smoothing	<i>PF</i> Regression	1.74 – 4.50	2.29
UNDP	Geographic Aggregation ²	P_2/F_2	1.96 – 4.70	2.48

¹Weighted by municipal populations of women 15–49 years old.

²For municipalities with fewer than 30,000 residents, the UNDP (2003) estimates municipal TFR as a multiple of the TFR for that municipality's *microregion* (the next-highest level of Census geography). The multiplier is a ratio of average reported parity in the municipality to average reported parity in the microregion. See text and Horta et al. (2005) for details.