ELSEVIER

Journal of Econometrics 113 (2003) 337–361

JOURNAL OF Econometrics

www.elsevier.com/locate/econbase

An econometric model of birth inputs and outputs for Native Americans

Kai Li[a,*], Dale J. Poirier[b]

[a] *Faculty of Commerce, University of British Columbia, 2053 Main Mall, Vancouver, BC, Canada V6T 1Z2*

[b] *Department of Economics, University of California, Irvine 3151 Social Sciences Plaza, Irvine, CA 92697-5100, USA*

Abstract

This paper presents a new model of the birth process of Native Americans with seven endogenous variables: four birth inputs *maternal smoking* (S), *drinking* (D), *prenatal care* (PC), and *weight gain* (WG), and three birth outputs *gestational age* (G), *birth length* (BL), and *birth weight* (BW). The model is a seven-equation simultaneous model with three endogenous dummies S, D, and PC. The data are taken from the National Longitudinal Survey of Youth (NLSY). We find that the four birth inputs are determined jointly and dependently among S, D, and PC, but independently of WG. S has negative systematic correlation with G. D and PC appear to have no sizeable systematic effect on G, BL, or BW. Except for the sizeable and positive correlation between the unexplained parts of S and G, there seem to be no unexplained common effects between the birth inputs and outputs. Moreover, G appears dependent on the exogenous size of the mother. BL is affected by the inputs mainly through WG. BW is affected by the inputs through their effects on G. Except for maternal weight, there is little correlation between the remaining exogenous variables and BW. Finally, the predictive density of BW for a typical pregnancy gives a mean weight of 3.240 kg.

JEL classification: C11; C34; I12; J13

Keywords: Bayesian; Birth length; Birth weight; Drinking; Gestation; Prenatal care; Simultaneity; Smoking; Weight gain

* Corresponding author. Tel.: +604-822-8353; fax: +604-822-4695.

E-mail addresses: kai.li@commerce.ubc.ca (K. Li), dpoirier@uci.edu (D.J. Poirier).

0304-4076/03/$ - see front matter
PII: S0304-4076(02)00206-3

1. Introduction

This paper is the first of a series by us (Li and Poirier, 2000, 2001a, b, 2002) analyzing *birth weight* (*BW*) and related birth outcomes. The model has its origins in Poirier (1998) and we extend the estimation procedures discussed in Chib and Greenberg (1998) and Li (1996, 1998).

BW is probably the single most important indicator of infant health (Institute of Health, 1985). It is also a significant predictor of infant mortality, morbidity, coronary heart disease, and learning disabilities later in life (Poirier, 1998). The vast majority of studies on BW, particularly in the biomedical literature are single-equation models that ignore the simultaneity issues, a notable exception is Permutt and Hebel (1989). An important contribution of this study is to address the simultaneity issue head on.

Our new model for the birth process is a nonlinear simultaneous equations model with the following features: (1) triangular coefficients of endogenous variables matrix, (2) mixed unlimited and limited (dichotomous) dependent variables, (3) zero restrictions on elements of both coefficients of endogenous variables matrix and coefficients of exogenous variables matrix, and (4) normal residuals. Existing work with a Bayesian orientation sharing some of these features are Chib and Greenberg (1998) and Li (1996, 1998). Chib and Greenberg (1998) concerns multivariate probit models. As in Chib and Greenberg (1998), diagonal elements of the variance–covariance matrix corresponding to dichotomous dependent variables are restricted to be unity and the Metropolis update (Chib and Greenberg, 1995) we take in this paper for each block of the variance–covariance matrix is essentially the same as Chib and Greenberg (1998). The model in Li (1996, 1998) has features of (1) and (4), and limited dependent variables (mixed dichotomous and censored).

The paper proceeds as follows. In Section 2 we introduce the data. In Section 3 we discuss our modeling framework and the prior distribution. We present computational details in Section 4 and report empirical results in Section 5. Some concluding remarks are offered in Section 6.

2. Data

The data we explore is the National Longitudinal Survey of Youth (NLSY). We focus on the experience of Native American women in the NLSY when they became mothers and on the birth of their children. We choose to analyze only *singleton first-born live births*, leaving aside sample selection problems arising from parity considerations (infants that are not first-born) and abortions. We drop births to Native American women in the military, births to Native American women no longer living in the U.S., and births with missing data. This leaves 81 singleton first-born live births to Native American women between 1979 and 1994 in the NLSY.

We work with seven endogenous variables: four birth inputs *maternal smoking* (S), *maternal drinking* (D), *first trimester prenatal care* (PC), and *maternal weight gain* (WG), and three birth outputs *gestational age* (G), *birth length* (BL), and *birth weight*

Table 1
Descriptive statistics: mean (std. dev. in mean) of endogenous variables

$z_1 = S$	$= 1$ if mother smoked during pregnancy $= 0$ otherwise	0.4198 (0.0552)
$z_2 = D$	$= 1$ if mother drank alcohol during pregnancy $= 0$ otherwise	0.4198 (0.0552)
$z_3 = \mathrm{PC}$	$= 1$ if prenatal care started in first trimester $= 0$ otherwise	0.8395 (0.0410)
$z_4 = \mathrm{WG}$	Weight gain net of BW in kg	12.06 (0.8188)
$z_5 = G$	Gestation in weeks	39.11 (0.2908)
$z_6 = \mathrm{BL}$	Birth length in cm	51.33 (0.3270)
$z_7 = \mathrm{BW}$	Birth weight in kg	3.347 (0.0751)

(BW). Our analysis conditions on 24 exogenous variables.[1] Our choice of exogenous (conditioning) variables is guided by the existing literature. Variable x_1 is the intercept term. Variables x_2–x_6 cover basic physical characteristics (the gender of the infant, and the age and size of the mother) which we expect to be very important in the birth output equations.[2] Variables x_7–x_{11} capture regional and temporal effects plus the intelligence and family income of the mother. We expect these variables to be important in the input equations, but hopefully not in the biologically based output equations. Variables x_{12}–x_{24} capture health insurance status, a variety of socioeconomic measures of the mother's family background, and four price indices. We expect these last 13 variables to be important only in the birth input equations. Means and standard deviations for all variables are given in Table 1 (endogenous variables denoted z) and Table 2 (exogenous variables denoted x).

3. Modeling

3.1. Modeling strategy

A priority for us is addressing simultaneity (i.e., endogeneity) of these birth input and output variables. The vast majority of studies on BW, particularly in the biomedical literature, are single-equation models that ignore simultaneity issues, except Permutt and Hebel (1989). When simultaneity issues are ignored, questions regarding the effects on endogenous variables z of changing exogenous variables x, assume an unresponsive mother who does not respond intelligently to changes in her environment. For example, suppose a component in x measures access to prenatal care. The meaningful answer to what is the effect on BW of changing this access requires a framework that allows

[1] The appendices in Li and Poirier (2000) contain a more detailed description of the variables involved.

[2] In this paper, we are not trying to explain fertility or the mother's pregnancy. Hence, variables like maternal age (x_3) are properly treated as exogenous in our analysis.

Table 2
Descriptive statistics: mean (std. dev. in mean) of exogenous variables

x_1	Constant	1.000 (0.0000)
x_2	Male child	0.4815 (0.0559)
x_3	Mother's age −23 yr.	−0.4198 (0.4749)
x_4	Body mass index (weight in kg/[height in m]^2) −24	0.3689 (0.5572)
x_5	Maternal height −162 (cm)	0.9363 (0.7211)
x_6	Maternal weight −63 (kg)	1.798 (1.591)
x_7	South	0.5926 (0.0549)
x_8	West	0.1852 (0.0434)
x_9	Calendar time −(19)85	−1.210 (0.4574)
x_{10}	(AFQT score/mean of NLSY women of same age) −1	0.0361 (0.0755)
x_{11}	Household income in $1000 −25	−3.682 (1.951)
x_{12}	No health insurance available	0.6914 (0.0516)
x_{13}	Missing health insurance availability	0.5309 (0.0558)
x_{14}	Number of adults in household −2	0.2716 (0.1069)
x_{15}	Number of quarters worked during pregnancy −3	−0.6296 (0.1614)
x_{16}	Number of maternal siblings −4	−0.3086 (0.2610)
x_{17}	Grandmother's education −12 yr.	−1.222 (0.2961)
x_{18}	Not being on time in school at age 14	0.1605 (0.0410)
x_{19}	Non-urban at age 14	0.2840 (0.0504)
x_{20}	No employed male in household at age 14	0.2469 (0.0482)
x_{21}	Cigarette price index	−0.0856 (0.0488)
x_{22}	Alcohol price index	−0.0502 (0.0203)
x_{23}	Medical services price index	−0.0839 (0.0394)
x_{24}	Food price index	−0.0488 (0.0200)

the mother to adjust the PC she employs. The standard BW regression, which contains measures of both the PC and access variables, is *not* designed to answer such questions.

While we draw upon the economics literature in our modeling, we do *not* invoke a formal optimization approach for the mother's decision making. We specify reduced form equations for the four inputs (see (1) below), and then a triangular specification (see (2)–(4) below) in which G (z_5) depends on the four inputs (z_1–z_4), and together BL (z_6) and BW (z_7) have a bivariate relationship depending on the four inputs and G. Our specification is overidentified and yields a fairly simple specification for all three output equations.

Following the strategy outlined in Poirier (1995, Chapter 10), we have chosen our initial model in the anticipation that a larger, more complicated one is *not* required. Of course there are many ways we could extend our initial model. One obvious way is to test some of the overidentifying restrictions that are a prerequisite for interpreting our model structurally. If our birth outcome production function reflects a biological transformation from birth inputs into birth outputs, then it should remain invariant over time (x_9) and not differ according to mothers' geographical region (x_7, x_8), AFQT score (x_{10}), or family income (x_{11}). Our prior in Section 3.3 reflects this viewpoint.

We now describe in detail our parametric view of the world.

3.2. An econometric window

Consider a sample of T independent singleton first-born live births indexed by the subscript i. Let $y_{i1}^* = [S_i^*, D_i^*, \mathrm{PC}_i^*]'$ $(i = 1, 2, \ldots, T)$ denote latent variables underlying the binary birth inputs $y_{i1} = [S_i, D_i, \mathrm{PC}_i]' = [\mathbf{1}(S_i^*), \mathbf{1}(D_i^*), \mathbf{1}(\mathrm{PC}_i^*)]'$ $(i = 1, 2, \ldots, T)$, where $\mathbf{1}(\bullet)$ denotes an indicator function which equals unity if the argument is positive and equals zero otherwise. For estimation, we partition the endogenous variables into inputs z_{i1} and outputs z_{i2} : $z_{i1}^* = [S_i^*, D_i^*, \mathrm{PC}_i^*, \mathrm{WG}_i]'$, $z_{i1} = [S_i, D_i, \mathrm{PC}_i, \mathrm{WG}_i]'$, $z_{i2} = [G_i, \mathrm{BL}_i, \mathrm{BW}_i]'$ $(i = 1, 2, \ldots, T)$. Stacking observations gives the data matrices: $Z_1^* = [z_{11}^*, z_{21}^*, \ldots, z_{T1}^*]'$, $Z_1 = [z_{11}, z_{21}, \ldots, z_{T1}]'$, $Z_2 = [z_{12}, z_{22}, \ldots, z_{T2}]'$, and $Z = [Z_1, Z_2]'$. Let x_i $(i = 1, 2, \ldots, T)$ denote $K \times 1$ $(K = 24)$ vectors of exogenous variables that are stacked into the $T \times K$ matrix $X = [x_1, x_2, \ldots, x_T]'$.

For birth i, suppose the four inputs are generated by

$$z_{i1}^* = \Delta_1' x_i + \varepsilon_{i1}, \tag{1}$$

where $\Delta_1 = [\Delta_S, \Delta_D, \Delta_{\mathrm{PC}}, \Delta_{\mathrm{WG}}]$ is an unrestricted $K \times 4$ matrix of unknown parameters. The three birth outputs are generated by

$$z_{i2}' \Gamma_2 = z_{i1}' \Gamma_1 + x_i' \Delta_2 + \varepsilon_{i2}', \tag{2}$$

where $\varepsilon_i = [\varepsilon_{i1}', \varepsilon_{i2}']' | x_i \sim$ i.i.d. $N_7(0_7, \Sigma)$ $(i = 1, 2, \ldots, T)$,

$$\Gamma_1 = \begin{bmatrix} \gamma_{S,G} & \gamma_{S,\mathrm{BL}} & \gamma_{S,\mathrm{BW}} \\ \gamma_{D,G} & \gamma_{D,\mathrm{BL}} & \gamma_{D,\mathrm{BW}} \\ \gamma_{\mathrm{PC},G} & \gamma_{\mathrm{PC},\mathrm{BL}} & \gamma_{\mathrm{PC},\mathrm{BW}} \\ \gamma_{\mathrm{WG},G} & \gamma_{\mathrm{WG},\mathrm{BL}} & \gamma_{\mathrm{WG},\mathrm{BW}} \end{bmatrix} = [\gamma_G \quad \gamma_{\mathrm{BL}} \quad \gamma_{\mathrm{BW}}], \tag{3}$$

$$\Gamma_2 = \begin{bmatrix} 1 & -\gamma_{G,\mathrm{BL}} & -\gamma_{G,\mathrm{BW}} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{4}$$

$$\Delta_2 = \begin{bmatrix} \delta_G & \delta_{\mathrm{BL}} & \delta_{\mathrm{BW}} \\ \delta_{5,G} & \delta_{5,\mathrm{BL}} & 0 \\ \delta_{6,G} & 0 & \delta_{6,\mathrm{BW}} \\ \Delta_{*,G} & \Delta_{*,\mathrm{BL}} & \Delta_{*,\mathrm{BW}} \\ 0_{13} & 0_{13} & 0_{13} \end{bmatrix}, \tag{5}$$

$\delta_G = [\delta_{1,G}, \delta_{2,G}, \delta_{3,G}, \delta_{4,G}]'$, $\delta_{\text{BL}} = [\delta_{1,\text{BL}}, \delta_{2,\text{BL}}, \delta_{3,\text{BL}}, \delta_{4,\text{BL}}]'$, $\delta_{\text{BW}} = [\delta_{1,\text{BW}}, \delta_{2,\text{BW}}, \delta_{3,\text{BW}}, \delta_{4,\text{BW}}]'$, $\Delta_{*,j} = [\delta_{7,j}, \ldots, \delta_{11,j}]'$ ($j = G$, BL, and BW), and

$$\Sigma = \left[\begin{array}{cccc|ccc} 1 & \sigma_{S,D} & \sigma_{S,\text{PC}} & \sigma_{S,\text{WG}} & \sigma_{S,G} & \sigma_{S,\text{BL}} & \sigma_{S,\text{BW}} \\ \sigma_{S,D} & 1 & \sigma_{D,\text{PC}} & \sigma_{D,\text{WG}} & \sigma_{D,G} & \sigma_{D,\text{BL}} & \sigma_{D,\text{BW}} \\ \sigma_{S,\text{PC}} & \sigma_{D,\text{WG}} & 1 & \sigma_{\text{PC},\text{WG}} & \sigma_{\text{PC},G} & \sigma_{\text{PC},\text{BL}} & \sigma_{\text{PC},\text{BW}} \\ \sigma_{S,\text{WG}} & \sigma_{D,\text{WG}} & \sigma_{\text{PC},\text{WG}} & \sigma^2_{\text{WG}} & \sigma_{\text{WG},G} & \sigma_{\text{WG},\text{BL}} & \sigma_{\text{WG},\text{BW}} \\ \hline \sigma_{S,G} & \sigma_{D,G} & \sigma_{\text{PC},G} & \sigma_{\text{WG},G} & \sigma^2_G & \sigma_{G,\text{BL}} & \sigma_{G,\text{BW}} \\ \sigma_{S,\text{BL}} & \sigma_{D,\text{BL}} & \sigma_{\text{PC},\text{BL}} & \sigma_{\text{WG},\text{BL}} & \sigma_{G,\text{BL}} & \sigma^2_{\text{BL}} & \sigma_{\text{BL},\text{BW}} \\ \sigma_{S,\text{BW}} & \sigma_{D,\text{BW}} & \sigma_{\text{PC},\text{BW}} & \sigma_{\text{WG},\text{BW}} & \sigma_{G,\text{BW}} & \sigma_{\text{BL},\text{BW}} & \sigma^2_{\text{BW}} \end{array}\right]$$

$$= \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}' & \Sigma_{22} \end{bmatrix}. \tag{6}$$

The specification in (1)–(6) reflects a view of the world that postulates a reduced form for the four inputs (S, D, PC, and WG), and a *triangular view* of the three outputs (G, BL, and BW) in which G is determined by the four inputs, and BL and BW are jointly determined as functions of the four inputs and G. The three output equations are identified by zero restrictions on maternal weight (x_6) in the BL equation, and on maternal height (x_5) in the BW equation. The model is not recursive because Σ is permitted to be nondiagonal. The model is nonlinear because of the jointly determined dummy endogenous variables (S, D, and PC). The specification of numerous zero restrictions on Δ_2 in (5) ensures that the order condition for identification is satisfied. We expect that we will *not* reject the maintained hypothesis $\Delta_* : \Delta_{*,G} = \Delta_{*,\text{BL}} = \Delta_{*,\text{BW}} = 0_5$ in favor of the alternative $\text{H}_\text{A} : \Delta_{*,G} \neq 0_5$ or $\Delta_{*,\text{BL}} \neq 0_5$ or $\Delta_{*,\text{BW}} \neq 0_5$.

Let θ denote the unique unknown elements in Γ_1, Γ_2, Δ, and Σ, and Θ denote the permissible parameter space. Let $\varphi_m(\bullet|\bullet,\bullet)$ denote an m-dimensional normal density with given mean vector and covariance. Because the conditional density of ε_{i2} given ε_{i1} is $f(\varepsilon_{i2}|\varepsilon_{i1},\Sigma) = \varphi_3(\varepsilon_{i2}|\Sigma_{12}'\Sigma_{11}^{-1}\varepsilon_{i1}, \Sigma_{2|1})$, where $\Sigma_{2|1} = \Sigma_{22} - \Sigma_{12}'\Sigma_{11}^{-1}\Sigma_{12}$, it follows using change-of-variable techniques, and noting from (2) that the Jacobian of the transformation from ε_{i2} to z_{i2} is unity due to the triangularity of Γ_2, the conditional distribution of the outputs z_{i2} given the inputs z_{i1}^* is

$$\begin{aligned} f(z_{i2}|z_{i1}^*, x_i, \theta) &= \varphi_3(\Gamma_2' z_{i2} - \Gamma_1' z_{i1} - \Delta_2' x_i | \Sigma_{12}'\Sigma_{11}^{-1}(z_{i1}^* - \Delta_1' x_i), \Sigma_{2|1}) \\ &= \varphi_3(z_{i2}|\mu_{i2} + \Sigma_{12}'\Sigma_{11}^{-1}(z_{i1}^* - \Delta_1' x_i), \Sigma_{2|1}), \end{aligned} \tag{7}$$

where under the maintained hypothesis H_*:

$$\mu_{i2} = \mu_{i2}[z_{i1}, G_i, x_i, \theta] = [0 \quad \gamma_{G,\text{BL}} G_i \quad \gamma_{G,\text{BW}} G_i]' + \Gamma_1' z_{i1} + \Delta_2' x_i = W_{i2}\beta_2, \tag{8}$$

W_{i2} is the 3×30 matrix

$$W_{i2} = \begin{bmatrix} z'_{i1} & x_i^{*\prime} & x_{i5} & x_{i6} & 0 & 0'_4 & 0'_4 & 0 & 0 & 0'_4 & 0'_4 & 0 \\ 0'_4 & 0'_4 & 0 & 0 & G_i & z'_{i1} & x_i^{*\prime} & x_{i5} & 0 & 0'_4 & 0'_4 & 0 \\ 0'_4 & 0'_4 & 0 & 0 & 0 & 0'_4 & 0_4 & 0 & G_i & z'_{i1} & x_i^{*\prime} & x_{i6} \end{bmatrix}, \tag{9}$$

$$x_i^* = [x_{i1},\ x_{i2},\ x_{i3},\ x_{i4}]',$$

and

$$\beta_2 = [\gamma'_G, \delta'_G, \delta_{5,G}, \delta_{6,G}, \gamma_{G,\mathrm{BL}}, \gamma'_{\mathrm{BL}}, \delta'_{\mathrm{BL}}, \delta_{5,\mathrm{BL}}, \gamma_{G,\mathrm{BW}}, \gamma'_{\mathrm{BW}}, \delta'_{\mathrm{BW}}, \delta_{6,\mathrm{BW}}]. \tag{10}$$

Combining the conditional density (7) with the marginal density $\varphi_4(z_{i1}^*|\Delta'_1, \Sigma_{11})$, and using properties of the multivariate normal density, it follows that the joint density of z_{i1}^* and z_{i2} is

$$\begin{aligned} f(z_{i1}^*, z_{i2}|x_i, \theta) &= \varphi_4(z_{i1}^*|\Delta'_1 x_i, \Sigma_{11})\varphi_3(z_{i2}|\mu_{i2} + \Sigma'_{12}\Sigma_{11}^{-1}(z_{i1}^* - \Delta'_1 x_i), \Sigma_{2|1}) \\ &= \phi_7\left(\begin{bmatrix} z_{i1}^* \\ z_{i2} \end{bmatrix} |\mu_i, \Sigma\right), \end{aligned} \tag{11}$$

where $\mu_i = W_i\beta$,

$$\beta = [\beta'_1, \beta'_2]', \qquad \beta_1 = \mathrm{vec}(\Delta_1) \tag{12}$$

and

$$W_i = \begin{bmatrix} I_4 \otimes x'_i & 0_{4\times 30} \\ 0_{3\times 96} & W_{i2} \end{bmatrix} \tag{13}$$

is 7×126. Note that (11) does *not* imply z_{i1}^* and z_{i2} are multivariate normal because μ_{i2} depends on elements in z_{i1}^* (through z_{i1}) and in z_{i2}. Under the alternative hypothesis $\mathrm{H_A}$, 15 additional columns are added to W_{i2}, and $\Delta_{*,j}$ ($j = G$, BL, and BW) are added to β_2.

Given the observed $y_{i1} = [S_i, D_i, \mathrm{PC}_i]'$, define the lower and upper integration limits

$$\underline{a}_{S_i} = \begin{cases} -\infty & \text{if } S_i = 0 \\ -x'_i\Delta_S & \text{if } S_i = 1 \end{cases}, \qquad \bar{a}_{S_i} = \begin{cases} -x'_i\Delta_S & \text{if } S_i = 0 \\ \infty & \text{if } S_i = 1 \end{cases}, \tag{14}$$

$$\underline{a}_{D_i} = \begin{cases} -\infty & \text{if } D_i = 0 \\ -x'_i\Delta_D & \text{if } D_i = 1 \end{cases}, \qquad \bar{a}_{D_i} = \begin{cases} -x'_i\Delta_D & \text{if } D_i = 0 \\ \infty & \text{if } D_i = 1 \end{cases}, \tag{15}$$

$$\underline{a}_{\mathrm{PC}_i} = \begin{cases} -\infty & \text{if } \mathrm{PC}_i = 0 \\ -x'_i\Delta_{\mathrm{PC}} & \text{if } \mathrm{PC}_i = 1 \end{cases}, \qquad \bar{a}_{\mathrm{PC}_i} = \begin{cases} -x'_i\Delta_{\mathrm{PC}} & \text{if } \mathrm{PC}_i = 0 \\ \infty & \text{if } \mathrm{PC}_i = 1 \end{cases}. \tag{16}$$

Then the joint density for all seven observed endogenous variables is

$$f(z_{i1},z_{i2}|x_i,\theta)=\int_{\underline{a}_{S_i}}^{\bar{a}_{S_i}}\int_{\underline{a}_{D_i}}^{\bar{a}_{D_i}}\int_{\underline{a}_{\mathrm{PC}_i}}^{\bar{a}_{\mathrm{PC}_i}} f(z_{i1}^*,z_{i2}|x_i,\theta)\,\mathrm{dPC}_i^*\,\mathrm{d}D_i^*\,\mathrm{d}S_i^*$$

$$=\int_{\underline{a}_{S_i}}^{\bar{a}_{S_i}}\int_{\underline{a}_{D_i}}^{\bar{a}_{D_i}}\int_{\underline{a}_{\mathrm{PC}_i}}^{\bar{a}_{\mathrm{PC}_i}} \varphi_4(z_{i1}^*|\Delta_1'x_i,\Sigma_{11})$$

$$+\varphi_3(z_{i2}|\mu_{i2}\Sigma_{12}'\Sigma_{11}^{-1}(z_{i1}^*-\Delta_1'x_i),\Sigma_{2|1}))\,\mathrm{dPC}_i^*\,\mathrm{d}D_i^*\,\mathrm{d}S_i^*. \tag{17}$$

Assuming independent sampling, we view the observed data, under H_*, through a $126+25$ (unknown elements in Σ) $=151$-dimensional parametric window given by likelihood

$$\mathscr{L}(\theta;Z,X)=\prod_{i=1}^{T} f(z_{i1},z_{i2}|x_i,\theta). \tag{18}$$

3.3. Our family of prior distributions

Our intention is to provide a *public prior*, i.e., a prior that captures other researchers' interests and permits them to reweight our Markov-chain Monte-Carlo (MCMC) simulations to obtain results corresponding to more tightly articulated prior beliefs (Geweke, 1999). Our prior is proper, but moderately diffuse. Our reading of the existing literature suggests the following broad properties will capture a bevy of researchers' professional opinions as well as ours.

We center our prior over a diagonal specification for Σ which is implicitly advocated in the biomedical literature that focuses on single-equation techniques. Due to the presence of three probit regressions in our system, the standard Wishart prior on the inverse of Σ is not appropriate and the natural conjugacy between the prior and likelihood breaks down. [3] For estimation, we partition the 25 unknown elements of Σ into eight blocks:

$$s_1=[\sigma_{S,D},\sigma_{S,\mathrm{PC}},\sigma_{D,\mathrm{PC}}]', \tag{19}$$

$$s_2=[\sigma_{S,\mathrm{WG}},\sigma_{S,G},\sigma_{S,\mathrm{BL}},\sigma_{S,\mathrm{BW}},\sigma_{D,\mathrm{WG}},\sigma_{D,G}]', \tag{20}$$

[3] An identification problem arises in our simultaneous equations model with multiple probits as any scale shift will not change the observed choices in the first three binary equations. One solution, which is implemented in this paper, is to only sample the identified elements in the variance–covariance matrix using the Metropolis–Hastings update suggested by Chib and Greenberg (1998). An alternative is to follow McCulloch and Rossi (1994) by adopting the standard Wishart prior for the variance–covariance matrix Σ. To reflect the fact that the vast majority of studies on birth outcomes are single-equation models that ignore simultaneity issues, we employ a fairly diffuse inverted Wishart prior with 20 degrees of freedom and centered over a diagonal matrix with diagonal elements: 1, 1, 1, 64, 4, 4, 0.01. Then we compute the full posterior over β and Σ and simply report the marginal posterior distribution of the identified parameters, e.g., Δ_s/σ_s and error correlation coefficients for the first three equations. We find posterior results from these two different priors on the variance–covariance matrix are not materially different.

$$s_3 = [\sigma_{D,\mathrm{BL}}, \sigma_{D,\mathrm{BW}}, \sigma_{\mathrm{PC,WG}}, \sigma_{\mathrm{PC},G}, \sigma_{\mathrm{PC,BL}}, \sigma_{\mathrm{PC,BW}}]', \tag{21}$$

$$s_4 = [\sigma_{\mathrm{WG},G}, \sigma_{\mathrm{WG,BL}}, \sigma_{\mathrm{WG,BW}}, \sigma_{G,\mathrm{BL}}, \sigma_{G,\mathrm{BW}}, \sigma_{\mathrm{BL,BW}}]', \tag{22}$$

$$s_5 = \sigma^2_{\mathrm{WG}}, \quad s_6 = \sigma^2_G, \quad s_7 = \sigma^2_{\mathrm{BL}}, \quad s_8 = \sigma^2_{\mathrm{BW}}. \tag{23}$$

The joint prior specification for the eight blocks is [using the notation of Poirier (1995, p. 111 (e)):

$$f(\Sigma) = \varphi_3(s_1|0_3, I_3)\left[\prod_{j=2}^{4}\varphi_6(s_j|0_6, 9I_6)\right] f_{IG}(s_5|2.5, 0.01) f_{IG}(s_6|2.5, 0.08) \times f_{IG}(s_7|2.5, 0.08) f_{IG}(s_8|2.5, 2), \tag{24}$$

and there is truncation of the above prior density to the set of positive definite matrices.

We put most effort into eliciting our prior beliefs about a mother "type" defined by the way we center the exogenous variables. The case in which all elements of x other than x_1 are zero describes the generic mother corresponding to the intercept in each equation. Our choice of centering is described in Table 2. For the reduced-form S^*, D^*, PC^*, and WG equations, the intercept corresponds to a north-central female birth in January 1985 to a 23-year-old mother, having access to health insurance, living with another adult, with a household income of $25,000, with body mass of 24 based on a height of 162 cm and a weight of 63 kg, who worked three of the four quarters in the year before giving birth, with four siblings, with mean AFQT score of other 23-year-old women in the NLSY, who was on-time in school (within one grade) in an urban household with an employed male at age 14, whose mother (the maternal grandmother) completed 12 years of education, and facing prices for cigarette, alcohol, medical services, and food at their 1984 levels. It corresponds to a mother for whom we would expect a favorable birth outcome. We express relatively informative priors for these intercepts.

Regarding the effects of birth inputs on birth outputs, there is substantial support that smoking (S) has negative consequences on birth outputs, particularly on gestation (G) and BW. The effect of drinking (D) on birth outputs is less clear and may even be positive, e.g., Meis et al. (1998). The effect of prenatal care (PC) on birth outputs is less obvious due to sample selection effects, e.g., Shiono and Behrman (1995), but we believe PC may be helpful for BW. We also believe weight gain (WG) and G have positive effects on birth length (BL) and BW. Our beliefs about the remaining effects of endogenous input variables on other endogenous output variables are fairly diffuse and centered over zero.

Among exogenous variables, we believe, *ceteris paribus*, male infants (x_2) are slightly longer and heavier than females; calendar time (x_9) has a slightly negative effect on S and D, a slightly positive effect on PC, and a very uncertain positive effect on WG. AFQT score (x_{10}) has a moderately negative effect on S and D, and a moderately positive effect on PC and WG. Grandmother's education (x_{17}) has a slightly

Table 3
Prior means (standard deviations) of Γ_1 and Γ_2 under both H_* and H_A: $\xi = 1$

Endogenous variable j	$\gamma_{j,G}$	$\gamma_{j,\mathrm{BL}}$	$\gamma_{j,\mathrm{BW}}$
S	-1.000	0.0000	-0.3500
	$(2\xi^{1/2})$	$(3\xi^{1/2})$	$(\xi^{1/2})$
D	0.0000	0.0000	0.0000
	$(2\xi^{1/2})$	$(3\xi^{1/2})$	$(\xi^{1/2})$
PC	0.0000	0.0000	0.1000
	$(2\xi^{1/2})$	$(3\xi^{1/2})$	$(\xi^{1/2})$
WG	0.0000	0.1000	0.1000
	$(2\xi^{1/2})$	$(3\xi^{1/2})$	$(\xi^{1/2})$
G	**−1.000**	0.0500	0.0100
	(0.0000)	$(\xi^{1/2})$	$(\xi^{1/2})$

negative effect on S and D, and a positive effect on PC and WG. Finally, not being on-time in school at age 14 (x_{18}) has a positive effect on S and D, and a negative effect on PC and WG. The effects of all other exogenous variables on the remaining endogenous variables are centered over zero with fairly wide standard errors.

For the unknown elements of Γ_1 in (3) and of Γ_2 in (4), we assume independent univariate normal priors with means and standard deviations given in Table 3, where ξ is a hyperparameter. We initially set $\xi = 1$, and then change it as part of a sensitivity analysis in Section 5.6.

The exogenous variables x_7–x_{24} play the role of instruments in birth output equations where they are subject to dogmatic zero restrictions satisfying the order condition for overidentification. Our priors reflect this instrument role. Specifically, we believe calendar time ($x_{9,j}$), AFQT score ($x_{10,j}$), grandmother's education ($x_{17,j}$), and not being on-time in school at age 14 ($x_{18,j}$) serve as particularly good instruments, so the nonzero prior means ($-0.2, -1, -0.5, 0.6$; or $0.2, 1, 0.5, -0.6$) for the coefficients of these variables imply substantial mass away from the point 0_4 at which the rank condition fails ($j = S, D, \mathrm{PC}$, and WG). The other variables among x_7–x_{24} also serve as instruments in subsequent equations, but we are less certain of their reliability as instruments, and so their prior means of zero fail the rank condition.

For the unknown elements in Δ, we assume normal priors with means and standard deviations given in Table 4, where ω_1 and ω_2 are hyperparameters initially set to $\omega_1 = 0.1936$ ($\omega_1^{1/2} = 0.44$) and $\omega_2 = 9$. For the mother described by the intercept, the prior probability of smoking is 0.50, and with a standard deviation $\omega_1^{1/2} = 0.44$ so that within one standard deviation of the mean the probability of smoking is between 0.33 and 0.67. Note the prior is structured so that the same probabilities hold across regions, regardless of the gender of the child or the availability of health insurance. Identical results are asserted for the priors of the D and PC equations.

All these beliefs are held *independently* across the unknown parameters in Γ_1, Γ_2, Δ, and Σ, except that beliefs across regions of the country which are exchangeable within each equation, i.e., the coefficients of these variables are equally correlated with covariance ω_1.

Table 4
Prior means (standard deviations) of Δ under H_*: $\omega_1 = 0.1936$, $\omega_2 = 9$

Variable		S	D	PC	WG	G	BL	BW
x_1	Intercept	0.0000	0.0000	0.0000	10.00	40.00	48.00	2.000
		$(\omega_1)^{1/2}$	$(\omega_1)^{1/2}$	$(\omega_1)^{1/2}$	$(6\omega_1^{1/2})$	$(4\omega_1^{1/2})$	$(4\omega_1^{1/2})$	$(2\omega_1^{1/2})$
x_2	Male child	0.0000	0.0000	0.0000	0.0000	0.0000	0.1000	0.1000
		$(\omega_1)^{1/2}$	$(\omega_1)^{1/2}$	$(\omega_1)^{1/2}$	$(6\omega_1^{1/2})$	$(4\omega_1^{1/2})$	$(4\omega_1^{1/2})$	$(2\omega_1^{1/2})$
x_3	Mother's age −23 yr.	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		$(\omega_2)^{1/2}$	$(\omega_2)^{1/2}$	$(\omega_2)^{1/2}$	$(6\omega_2^{1/2})$	$(4\omega_2^{1/2})$	$(4\omega_2^{1/2})$	$(2\omega_2^{1/2})$
x_4	Body mass index −24	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		$(\omega_2)^{1/2}$	$(\omega_2)^{1/2}$	$(\omega_2)^{1/2}$	$(6\omega_2^{1/2})$	$(4\omega_2^{1/2})$	$(4\omega_2^{1/2})$	$(2\omega_2^{1/2})$
x_5	Maternal height −162 cm	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	**0.0000**
		$(\omega_2)^{1/2}$	$(\omega_2)^{1/2}$	$(\omega_2)^{1/2}$	$(6\omega_2^{1/2})$	$(4\omega_2^{1/2})$	$(4\omega_2^{1/2})$	**(0.0000)**
x_6	Maternal weight −63 kg	0.0000	0.0000	0.0000	0.0000	0.0000	**0.0000**	0.0000
		$(\omega_2)^{1/2}$	$(\omega_2)^{1/2}$	$(\omega_2)^{1/2}$	$(6\omega_2^{1/2})$	$(4\omega_2^{1/2})$	**(0.0000)**	$(2\omega_2^{1/2})$
x_7	South	0.0000	0.0000	0.0000	0.0000	**0.0000**	**0.0000**	**0.0000**
		$(2\omega_1)^{1/2}$	$(2\omega_1)^{1/2}$	$(2\omega_1)^{1/2}$	$(6[2\omega_1]^{1/2})$	**(0.0000)**	**(0.0000)**	**(0.0000)**
x_8	West	0.0000	0.0000	0.0000	0.0000	**0.0000**	**0.0000**	**0.0000**
		$(2\omega_1)^{1/2}$	$(2\omega_1)^{1/2}$	$(2\omega_1)^{1/2}$	$(6[2\omega_1]^{1/2})$	**(0.0000)**	**(0.0000)**	**(0.0000)**
x_9	Calendar Time −(19)85	−0.2000	−0.2000	0.2000	0.2000	**0.0000**	**0.0000**	**0.0000**
		$(\omega_2)^{1/2}$	$(\omega_2)^{1/2}$	$(\omega_2)^{1/2}$	$(6\omega_2^{1/2})$	**(0.0000)**	**(0.0000)**	**(0.0000)**
x_{10}	(AFQT score/mean of same age) −1	−1.000	−1.000	1.000	1.000	**0.0000**	**0.0000**	**0.0000**
		$(\omega_1)^{1/2}$	$(\omega_1)^{1/2}$	$(\omega_1)^{1/2}$	$(6\omega_1^{1/2})$	**(0.0000)**	**(0.0000)**	**(0.0000)**
x_{11}	Household income in \$1000 −25	0.0000	0.0000	0.0000	0.0000	**0.0000**	**0.0000**	**0.0000**
		$(\omega_2)^{1/2}$	$(\omega_2)^{1/2}$	$(\omega_2)^{1/2}$	$(6\omega_2^{1/2})$	**(0.0000)**	**(0.0000)**	**(0.0000)**
x_{12}	No health insurance available	0.0000	0.0000	0.0000	0.0000	**0.0000**	**0.0000**	**0.0000**
		$(\omega_1)^{1/2}$	$(\omega_1)^{1/2}$	$(\omega_1)^{1/2}$	$(6\omega_1^{1/2})$	**(0.0000)**	**(0.0000)**	**(0.0000)**
x_{13}	Missing health insurance availability	0.0000	0.0000	0.0000	0.0000	**0.0000**	**0.0000**	**0.0000**
		$(2\omega_1)^{1/2}$	$(2\omega_1)^{1/2}$	$(2\omega_1)^{1/2}$	$(6[2\omega_1]^{1/2})$	**(0.0000)**	**(0.0000)**	**(0.0000)**
x_{14}	Number of adults in household −2	0.0000	0.0000	0.0000	0.0000	**0.0000**	**0.0000**	**0.0000**
		$(\omega_2)^{1/2}$	$(\omega_2)^{1/2}$	$(\omega_2)^{1/2}$	$(6\omega_2^{1/2})$	**(0.0000)**	**(0.0000)**	**(0.0000)**
x_{15}	No. of quarters worked during pregnancy −3	0.0000	0.0000	0.0000	0.0000	**0.0000**	**0.0000**	**0.0000**
		$(\omega_2)^{1/2}$	$(\omega_2)^{1/2}$	$(\omega_2)^{1/2}$	$(6\omega_2^{1/2})$	**(0.0000)**	**(0.0000)**	**(0.0000)**

Table 4 (*continued*)

Variable		*S*	*D*	PC	WG	*G*	BL	BW
x_{16}	Number of maternal siblings −4	0.0000	0.0000	0.0000	0.0000	**0.0000**	**0.0000**	**0.0000**
		$(\omega_2)^{1/2}$	$(\omega_2)^{1/2}$	$(\omega_2)^{1/2}$	$(6\omega_2^{1/2})$	**(0.0000)**	**(0.0000)**	**(0.0000)**
x_{17}	Grandmother's education −12 yr.	−0.5000	−0.5000	0.5000	0.5000	**0.0000**	**0.0000**	**0.0000**
		$(\omega_1)^{1/2}$	$(\omega_1)^{1/2}$	$(\omega_1)^{1/2}$	$(6\omega_1^{1/2})$	**(0.0000)**	**(0.0000)**	**(0.0000)**
x_{18}	Not being on time in school at age 14	0.6000	0.6000	−0.6000	−0.6000	**0.0000**	**0.0000**	**0.0000**
		$(\omega_1)^{1/2}$	$(\omega_1)^{1/2}$	$(\omega_1)^{1/2}$	$(6\omega_1^{1/2})$	**(0.0000)**	**(0.0000)**	**(0.0000)**
x_{19}	Non-urban at age 14	0.0000	0.0000	0.0000	0.0000	**0.0000**	**0.0000**	**0.0000**
		$(\omega_1)^{1/2}$	$(\omega_1)^{1/2}$	$(\omega_1)^{1/2}$	$(6\omega_1^{1/2})$	**(0.0000)**	**(0.0000)**	**(0.0000)**
x_{20}	No employed male in household at age 14	0.0000	0.0000	0.0000	0.0000	**0.0000**	**0.0000**	**0.0000**
		$(\omega_1)^{1/2}$	$(\omega_1)^{1/2}$	$(\omega_1)^{1/2}$	$(6\omega_1^{1/2})$	**(0.0000)**	**(0.0000)**	**(0.0000)**
x_{21}	Cigarette price index	0.0000	0.0000	0.0000	0.0000	**0.0000**	**0.0000**	**0.0000**
		$(\omega_2)^{1/2}$	$(\omega_2)^{1/2}$	$(\omega_2)^{1/2}$	$(6\omega_2^{1/2})$	**(0.0000)**	**(0.0000)**	**(0.0000)**
x_{22}	Alcohol price index	0.0000	0.0000	0.0000	0.0000	**0.0000**	**0.0000**	**0.0000**
		$(\omega_2)^{1/2}$	$(\omega_2)^{1/2}$	$(\omega_2)^{1/2}$	$(6\omega_2^{1/2})$	**(0.0000)**	**(0.0000)**	**(0.0000)**
x_{23}	Medical services price index	0.0000	0.0000	0.0000	0.0000	**0.0000**	**0.0000**	**0.0000**
		$(\omega_2)^{1/2}$	$(\omega_2)^{1/2}$	$(\omega_2)^{1/2}$	$(6\omega_2^{1/2})$	**(0.0000)**	**(0.0000)**	**(0.0000)**
x_{24}	Food price index	0.0000	0.0000	0.0000	0.0000	**0.0000**	**0.0000**	**0.0000**
		$(\omega_2)^{1/2}$	$(\omega_2)^{1/2}$	$(\omega_2)^{1/2}$	$(6\omega_2^{1/2})$	**(0.0000)**	**(0.0000)**	**(0.0000)**

As noted in Section 3.2, we choose a highly overidentified specification for our maintained hypothesis H_*, and a less restricted specification H_A as an alternate hypothesis that we *expect* will not lead to rejecting H_*. Under both H_* and H_A, our priors for all parameters in the birth input equations are the same.

4. Computation

Given that there are three Probit equations in our simultaneous equations model, it is computationally intensive to evaluate the likelihood function (18), so we work with the joint posterior distribution of both the parameters and the latent data Y_1^* conditional on the observed data. Some earlier applications of this technique can be found in Chib (1992) on Tobit models and Albert and Chib (1993) on binary and polychotomous response data. From Bayes theorem, the augmented posterior is

$$f(\beta, \Sigma, Y_1^* \mid Z, X) \propto f(\beta, \Sigma) f(Y_1^* \mid \beta, \Sigma, Z, X) \mathscr{L}(\beta, \Sigma; Z, X). \tag{25}$$

The Markov-chain sampling scheme can be constructed from the distributions $f(Y_1^*|\beta, \Sigma, Z, X), f(\beta|\Sigma, Y_1^*, Z, X)$, and $f(\Sigma|\beta, Y_1^*, Z, X)$. More specifically, in the first block, we sample the latent data Y_1^* by applying the method developed in Geweke (1991); in the second block, the conditional distribution of β is multivariate normal, from which it is easy to draw variates directly; and finally, in the third block, we sample elements of the variance–covariance matrix Σ using the Metropolis update, see Chib and Greenberg (1995) for an overview.

Let $q(\Sigma^\dagger|\Sigma, \beta, Y_1^*, Z, X)$ denote a proposal density that generates candidate values $\Sigma^\dagger$ given the current value Σ. The choice of the proposal density is given later. The Metropolis update works in the following two steps:

(i) Sample a draw $\Sigma^\dagger$ given Σ from the proposal density $q(\Sigma^\dagger|\Sigma, \beta, Y_1^*, Z, X)$.
(ii) Move to $\Sigma^\dagger$ with probability

$$\mathrm{p}(\Sigma, \Sigma^\dagger) = \min\left(\frac{f(\Sigma^\dagger) f(Y_1^*|\beta, \Sigma^\dagger, Z, X) \mathscr{L}(\beta, \Sigma^\dagger; Z, X) 1(\Sigma^\dagger \in C) q(\Sigma|\Sigma^\dagger, \beta, Y_1^*, Z, X)}{f(\Sigma) f(Y_1^*|\beta, \Sigma, Z, X) \mathscr{L}(\beta, \Sigma; Z, X) 1(\Sigma \in C) q(\Sigma^\dagger|\Sigma, \beta, Y_1^*, Z, X)}, 1\right), \tag{26}$$

and stay at Σ with probability $1 - \mathrm{p}(\Sigma, \Sigma^\dagger)$.

Note that because the positive definiteness constraint on the variance–covariance matrix $\mathbf{1}(\Sigma^\dagger \in \mathrm{C})$ is part of the target density, the proposal density q is not truncated to C. Thus, if $\Sigma^\dagger$ is not positive definite, the conditional posterior is zero, and the proposal value is rejected with certainty, i.e., without having to evaluate the prior and the likelihood, we stay at state Σ.

In our empirical application, there are 25 unknown elements in Σ. Thus, it can be a challenging task to search for a suitable candidate-generating density. Following Chib and Greenberg (1998), we partition Σ into blocks (19)–(23) and adopt the random walk chain to generate proposal values in sequence through the various blocks. In particular, we use a multivariate normal candidate-generating density for the off-diagonal blocks of Σ (19)–(22), and a univariate normal density for the diagonal elements of Σ (23). The mean of the normal is given by the previous draw and the variance is calibrated so that the acceptance probability is reasonable. See Chib and Greenberg (1995) for some rough guidelines on choices of the variance matrices used in the random walk chains, and Li and Poirier (2000) for further implementation details.

To obtain the posterior estimates of model parameters, we take a run of 100,000 replications from our MCMC algorithm and discard the initial 20,000 to mitigate the startup effect. Preliminary runs are used to calibrate the variance matrices for our normal candidate generating densities used in the Metropolis update. We use a plot of all MCMC draws and the Convergence Diagnostic and Output Analysis software (Cowles and Carlin, 1996) to ensure convergence.

In the paper, we also consider the question of comparing different simultaneous equations models of the birth inputs and outputs, for example, one benchmark specification is that there is no simultaneity among the birth inputs and outputs and Σ is diagonal. These alternative models may be compared by Bayes factors, or ratios of model marginal likelihoods. We compute the marginal likelihood of each model specification using the method developed in Chib and Jeliazkov (2001).

Finally, we compute the population R^2 for each birth output equation as Li and Poirier (2001a). The measure is developed by Carter and Nagar (1977). It has the same interpretation as the familiar coefficient of determination used with the classical linear model, and at the same time it explicitly includes all the restrictions that serve to identify the structural equation.

5. Empirical results

5.1. Evidence of structure

The Bayes factor in favor of our maintained specification H_* versus the more general H_A is an overwhelming e^{48}. The five additional variables south, west, calendar time, AFQT score, and household income (x_7–x_{11}) add little to the G equation (like their other exogenous counterparts). There is some regional variation in both the BL and BW equations, and temporal and income effects in the BL equation. But the overall effect is negligible as indicated by the Bayes factor. Since our prior odds also favor H_*, we feel reassured in adopting our maintained specification, which is effectively the same as model averaging for all reasonable prior odds ratios.

Given our 151-dimensional window there are ample opportunities for pretesting, but we do not engage in it (an exception is the diagnostic testing of H_*). Instead we report posterior means and standard deviations. To give a quick, visual indication of posterior mass around the means, we indicate the relative size of the posterior mean to the posterior standard deviation by the border of table cells as described in Table 5. We first report results for our baseline hyperparameter specification, $\xi = 1$, $\omega_1 = 0.1936$, $\omega_2 = 9$, leaving to Section 5.6 a discussion of the sensitivity of our results.

Table 5
Notational conventions in subsequent tables

□	Absolute value of mean between one and two standard deviations
⧈	Absolute value of mean between two and three standard deviations
▣	Absolute value of mean more than three standard deviations
bold	Standard deviation equal to zero

Table 6
Posterior means (standard deviations) of across-equation correlations and variances in Σ: $\xi = 1$, $\omega_1 = 0.1936$, $\omega_2 = 9$

	S^*	D^*	PC*	WG	G	BL	BW
S^*	**1.000** (**0.0000**)	⧈ 0.3118 (0.1546)	□ −0.2480 (0.1774)	0.0066 (0.0925)	□ 0.2446 (0.2377)	−0.2017 (0.2060)	−0.2352 (0.2557)
D^*		**1.000** (**0.0000**)	□ −0.3522 (0.1846)	−0.0246 (0.0895)	0.0653 (0.1898)	0.0789 (0.2310)	−0.1468 (0.2674)
PC*			**1.000** (**0.0000**)	0.0500 (0.1011)	0.1078 (0.2147)	−0.0451 (0.2270)	0.0722 (0.2548)
WG				▣ 64.07 (11.44)	−0.0028 (0.0447)	−0.0012 (0.0437)	0.0755 (0.1736)
G					▣ 7.993 (1.372)	□ 0.1340 (0.0858)	□ 0.1681 (0.1351)
BL						▣ 8.652 (1.387)	▣ 0.4148 (0.0998)
BW							▣ 0.4367 (0.0832)

Note: Variances for S^*, D^* and PC* are normalized to unity. Off-diagonal elements are given as *correlations*, not covariances.

5.2. System considerations

Posterior means and standard deviations of the correlations and variances in Σ are given in Table 6. The largest correlation is between the disturbances in the BL and BW equations with posterior mean 0.4148. There is moderate correlation among the three

Table 7
Posterior means (standard deviations) of Γ_1 and Γ_2: $\xi = 1,\ \omega_1 = 0.1936,\ \omega_2 = 9$

Endogenous variable j	$\gamma_{j,G}$	$\gamma_{j,\mathrm{BL}}$	$\gamma_{j,\mathrm{BW}}$
S	−0.1967 (1.094)	−0.1968 (1.086)	−0.1136 (0.2908)
D	−0.7112 (0.9714)	−0.8107 (1.202)	0.0268 (0.3135)
PC	−0.3798 (1.101)	0.1126 (1.319)	−0.0015 (0.3246)
WG	0.0061 (0.0430)	0.1026 (0.0456)	0.0036 0.0167)
G	**−1.000** (**0.0000**)	0.0726 (0.0555)	0.0362 (0.0231)

discrete input disturbances, but not with WG, and among the three output disturbances. Besides the correlation between the disturbances in the BL and BW output equations, there is minor correlation across the G and BL output equations. But most notably, there is *little* correlation between input and output disturbances. This is consistent with our prior specification on Σ. The overall weak input–output correlations bring into question whether simultaneous equation techniques are necessary. [4] The relatively small sample size compared to the 151 unknown parameters and the loosely parameterized input equations probably account for the weak association between birth inputs and outputs.

5.3. Endogenous effects

The systematic endogenous effects are captured by Γ_1 and Γ_2. Posterior means and standard deviations are given in Table 7. The direct effects of the three discrete inputs (S, D, and PC) on the outputs are particularly small. WG seems to only directly affect BL. G is basically unexplained. While this may be due to measurement error problems with the G data, G does have a small positive systematic effect on BL and BW. [5] The posterior mean of $\gamma_{G,\mathrm{BW}}$ suggests another week of gestation results in a 36 g increase in BW.

[4] The Bayes factor favors the single-equation model over the simultaneous equations model. Then when comparing the simultaneous equations model with a model where there is zero correlation between birth inputs and outputs (i.e., the variance–covariance matrix Σ is block diagonal). The data prefers our simultaneous equations model.

[5] The G data exhibit a markedly large spike at $G = 39$, with very small adjacent spikes and then moderately large spikes further away. Our treating of G as continuous and free of measurement error is an assumption we hope to weaken in subsequent analysis.

5.4. Exogenous effects

Posterior means and standard deviations for the effects of exogenous variables in each equation, as captured by Δ, are reported in Table 8.

Consider the effects of the instruments x_7–x_{24}. The validity of the instruments west (x_8), no health insurance available (x_{12}), number of maternal siblings (x_{17}), grandmother's education (x_{17}), no employed males in household at age 14 (x_{20}), and three price indices (x_{22}–x_{24}) is in doubt. The coefficients of other instruments, however, all indicate mass away from zero in at least one input equation.

Next consider x_2–x_6. Surprisingly, male child (x_2) does not have much impact on BL and BW, but it does appear important in the S and PC equations. Maternal height (x_5) and weight (x_6) seem less important in the output equations than in the input equations. But a more careful analysis takes into account the exogenous effects of maternal height (x_5) and weight (x_6) through BMI (x_4). Table 9 reports the posterior means and standard deviations of the partial derivatives of the exogenous variable effects of maternal height and weight in each equation. These results confirm the earlier observations. The effect of maternal weight on BW seems especially important.

Explaining WG seems particularly difficult, and interestingly, it is the birth input that shows the greatest impact in the BL equation (recall Table 7). The posterior means of the effects of exogenous variables on WG are negligible and often of surprising sign.

Finally, the R^2's for the G, BL, and BW equations are 0.1202, 0.1141, and 0.2035, respectively.

5.5. Prediction

Given out-of-sample values of the exogenous variables $\tilde{x}$, the predictive density for the out of sample of $\tilde{z}^* = [\tilde{z}_1^*, \tilde{z}_2]' = [\tilde{S}^*, \tilde{D}^*, \tilde{\text{PC}}^*, \tilde{\text{WG}}, \tilde{G}, \tilde{\text{BL}}, \tilde{\text{BW}}]'$ is

$$\begin{aligned} f(\tilde{z}_1^*, \tilde{z}_2|\tilde{x}, Z, X) &= \int_\Theta f(\tilde{z}_1^*, \tilde{z}_2|\tilde{x}, \theta) f(\theta|Z,X)\,\mathrm{d}\theta \\ &= \int_\Theta \varphi_4(\tilde{z}_1^*|\Delta_1'\tilde{x}, \Sigma_{11})\varphi_3(\tilde{z}_2|\tilde{\mu}_2 + \Sigma_{12}'\Sigma_{11}^{-1}(\tilde{z}_1^* - \Delta_1'\tilde{x}), \Sigma_{2|1}) \\ &\quad f(\theta|Z,X)\,\mathrm{d}\theta, \end{aligned} \tag{27}$$

where $\tilde{\mu}_2 = \tilde{\mu}_2(\tilde{z}_1, \tilde{G}, \tilde{x}, \beta) = \tilde{W}_2\beta$,

$$\tilde{W}_2 = \begin{bmatrix} \tilde{z}_1' & \tilde{x}^{*\prime} & \tilde{x}_5 & \tilde{x}_6 & 0 & 0_4' & 0_4' & 0 & 0 & 0_4' & 0_4' & 0 \\ 0_4' & 0_4' & 0 & 0 & \tilde{G} & \tilde{z}_1' & \tilde{x}^{*\prime} & \tilde{x}_5 & 0 & 0_4' & 0_4' & 0 \\ 0_4' & 0_4' & 0 & 0 & 0 & 0_4' & 0_4' & 0 & \tilde{G} & \tilde{z}_1' & \tilde{x}^{*\prime} & \tilde{x}_6 \end{bmatrix} \tag{28}$$

and $\tilde{z}_1 = [\tilde{S}, \tilde{D}, \tilde{\text{PC}}, \tilde{\text{WG}}]' = [\mathbf{1}(\tilde{y}_1^*), \mathbf{1}(\tilde{y}_2^*), \mathbf{1}(\tilde{y}_3^*), \tilde{\text{WG}}]'$. We will concentrate on the predictive distribution for BW, obtained from (27) by integrating-out inputs and other

Table 8
Posterior means (standard deviations) of Δ: $\xi = 1$, $\omega_1 = 0.1936$, $\omega_2 = 9$

	Variable	S	D	PC	WG	G	BL	BW
x_1	Intercept	0.1204	0.0833	0.3478	9.990	39.55	47.35	1.857
		(0.3035)	(0.3080)	(0.3435)	(1.851)	(1.055)	(1.705)	(0.8496)
x_2	Male child	0.3487	−0.0095	0.7742	−0.1443	0.4540	0.4028	0.0765
		(0.2487)	(0.2449)	(0.2876)	(1.489)	(0.6229)	(0.6407)	(0.1531)
x_3	Mother's age −23 yr.	−0.0736	0.2202	0.1547	0.2917	−0.0178	−0.0429	−0.0066
		(0.0848)	(0.0924)	(0.1078)	(0.5083)	(0.0844)	(0.0893)	(0.0209)
x_4	Body mass index −24	0.1911	0.8289	0.9296	1.976	−0.7709	0.0625	−0.0005
		(0.4783)	(0.4649)	(0.4853)	(2.638)	(0.8177)	(0.0706)	(0.0409)
x_5	Maternal height −162 cm	0.0613	0.2653	0.2531	0.5451	−0.3070	0.0387	**0.0000**
		(0.1481)	(0.1465)	(0.1542)	(0.8290)	(0.2579)	(0.0532)	(**0.0000**)
x_6	Maternal weight −63 kg	−0.0796	−0.3246	−0.3544	−0.6503	0.2946	**0.0000**	0.0191
		(0.1804)	(0.1750)	(0.1850)	(0.9914)	(0.3075)	(**0.0000**)	(0.0144)
x_7	South	−0.4757	−0.1877	−0.1193	1.792	**0.0000**	**0.0000**	**0.0000**
		(0.3275)	(0.3337)	(0.3915)	(2.008)	(**0.0000**)	(**0.0000**)	(**0.0000**)
x_8	West	0.2280	0.0602	0.1241	−0.3904	**0.0000**	**0.0000**	**0.0000**
		(0.3929)	(0.4206)	(0.4813)	(2.409)	(**0.0000**)	(**0.0000**)	(**0.0000**)

Table 8 (*continued*)

	Variable	*S*	*D*	PC	WG	*G*	BL	BW
x_9	Calendar time −(19)85	0.3720	−0.2141	−0.5477	−0.1864	**0.0000**	**0.0000**	**0.0000**
		(0.3004)	(0.2963)	(0.3476)	(1.836)	**(0.0000)**	**(0.0000)**	**(0.0000)**
x_{10}	(AFQT score/mean of same age) −1	−0.5711	−0.2343	0.2118	0.4115	**0.0000**	**0.0000**	**0.0000**
		(0.2575)	(0.2709)	(0.3004)	(1.501)	**(0.0000)**	**(0.0000)**	**(0.0000)**
x_{11}	Household income in $1000 −25	−0.0077	0.0158	−0.0034	−0.0524	**0.0000**	**0.0000**	**0.0000**
		(0.0130)	(0.0138)	(0.0142)	(0.0765)	**(0.0000)**	**(0.0000)**	**(0.0000)**
x_{12}	No health insurance available	−0.0442	0.0090	0.1658	−1.106	**0.0000**	**0.0000**	**0.0000**
		(0.2949)	(0.2956)	(0.3272)	(1.783)	**(0.0000)**	**(0.0000)**	**(0.0000)**
x_{13}	Missing health insurance availability	−0.4918	−0.5810	0.0944	2.462	**0.0000**	**0.0000**	**0.0000**
		(0.3892)	(0.3966)	(0.4338)	(2.271)	**(0.0000)**	**(0.0000)**	**(0.0000)**
x_{14}	Number of adults in household −2	−0.3236	0.1267	0.5260	−0.1144	**0.0000**	**0.0000**	**0.0000**
		(0.2002)	(0.1812)	(0.2928)	(1.010)	**(0.0000)**	**(0.0000)**	**(0.0000)**
x_{15}	No. of quarters worked during pregnancy −3	0.0207	−0.3350	−0.2693	−0.9104	**0.0000**	**0.0000**	**0.0000**
		(0.1632)	(0.1662)	(0.2040)	(0.9445)	**(0.0000)**	**(0.0000)**	**(0.0000)**
x_{16}	Number of maternal siblings −4	−0.0004	0.0784	−0.0009	0.1914	**0.0000**	**0.0000**	**0.0000**
		(0.0744)	(0.0817)	(0.0979)	(0.4590)	**(0.0000)**	**(0.0000)**	**(0.0000)**
x_{17}	Grandmother's education −12 yr.	−0.0608	0.0543	0.0327	0.4149	**0.0000**	**0.0000**	**0.0000**
		(0.0757)	(0.0844)	(0.0901)	(0.4874)	**(0.0000)**	**(0.0000)**	**(0.0000)**

Table 8 (*continued*)

	Variable	S	D	PC	WG	G	BL	BW
x_{18}	Not being on time in school at age 14	0.6280	0.1195	−0.4984	0.0765	**0.0000**	**0.0000**	**0.0000**
		(0.3408)	(0.3254)	(0.3508)	(1.959)	(**0.0000**)	(**0.0000**)	(**0.0000**)
x_{19}	Non-urban at age 14	0.3954	−0.2202	0.2683	−0.9039	**0.0000**	**0.0000**	**0.0000**
		(0.2784)	(0.2886)	(0.3259)	(1.721)	(**0.0000**)	(**0.0000**)	(**0.0000**)
x_{20}	No employed male in household at age 14	0.0428	−0.2581	0.0500	0.6362	**0.0000**	**0.0000**	**0.0000**
		(0.2978)	(0.3102)	(0.3260)	(1.789)	(**0.0000**)	(**0.0000**)	(**0.0000**)
x_{21}	Cigarette price index	−2.445	0.7501	1.892	3.818	**0.0000**	**0.0000**	**0.0000**
		(1.870)	(1.816)	(2.052)	(10.95)	(**0.0000**)	(**0.0000**)	(**0.0000**)
x_{22}	Alcohol price index	1.731	0.2797	1.816	−7.113	**0.0000**	**0.0000**	**0.0000**
		(2.743)	(2.681)	(2.798)	(16.68)	(**0.0000**)	(**0.0000**)	(**0.0000**)
x_{23}	Medical services price index	−0.4985	−0.9403	1.697	−3.191	**0.0000**	**0.0000**	**0.0000**
		(2.652)	(2.730)	(2.750)	(16.27)	(**0.0000**)	(**0.0000**)	(**0.0000**)
x_{24}	Food price index	−1.435	1.045	1.016	2.494	**0.0000**	**0.0000**	**0.0000**
		(2.764)	(2.787)	(2.794)	(16.39)	(**0.0000**)	(**0.0000**)	(**0.0000**)

Table 9
Exogenous effects of maternal height and weight: $\xi = 1$, $\omega_1 = 0.1936$, $\omega_2 = 9$

Equation	Effect of maternal height	Effect of maternal weight
S	0.0018 (0.0112)	−0.0026 (0.0052)
D	0.0075 (0.0116)	−0.0033 (0.0054)
PC	−0.0080 (0.0130)	−0.0001 (0.0061)
WG	−0.0404 (0.1742)	0.1027 (0.0823)
G	−0.0786 (0.0556)	0.0008 (0.0264)
BL	0.0202 (0.0569)	0.0238 (0.0269)
BW	0.0001 (0.0121)	0.0189 (0.0062)

outputs:

$$f(\tilde{\mathrm{BW}}|\tilde{x},Z,X) = \int_{\Theta}\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{\mathfrak{R}^4} f(\tilde{z}_1^*,\tilde{z}_2|\theta)\,\mathrm{d}\tilde{z}_1^*\,\mathrm{d}\tilde{G}\,\mathrm{d}\tilde{\mathrm{BL}}\right] f(\theta|Z,X)\,\mathrm{d}\theta$$
$$= \int_{\Theta}\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{\mathfrak{R}^4} \varphi_4(\tilde{z}_1^*|\Delta_1'\tilde{x},\Sigma_{11})\varphi_3(\tilde{z}_2|\tilde{\mu}_2\right.$$
$$\left. + \Sigma_{12}'\Sigma_{11}^{-1}(\tilde{z}_1^* - \Delta_1'\tilde{x}),\Sigma_{2|1})\,\mathrm{d}\tilde{z}_1^*\,\mathrm{d}\tilde{G}\,\mathrm{d}\tilde{\mathrm{BL}}\right] f(\theta|Z,X)\,\mathrm{d}\theta. \qquad (29)$$

Note that the integral in (29) with respect to $\tilde{\mathrm{BL}}$ can be done analytically. Evaluating (29) for the reference mother corresponding to the intercept yields the predictive density shown in Fig. 1 with a predictive mean and standard deviation of 3.240 and 0.7115, respectively. Fig. 1 also depicts the very diffuse prior predictive density embodying only the informative prior and no data.

5.6. Sensitivity analysis

Our sensitivity analysis divides by four our baseline hyperparameter values $\xi = 1$, $\omega_1 = 0.1936$, $\omega_2 = 9$ to obtain $\xi = 0.25$, $\omega_1 = 0.0484$, $\omega_2 = 2.25$ and multiplies by

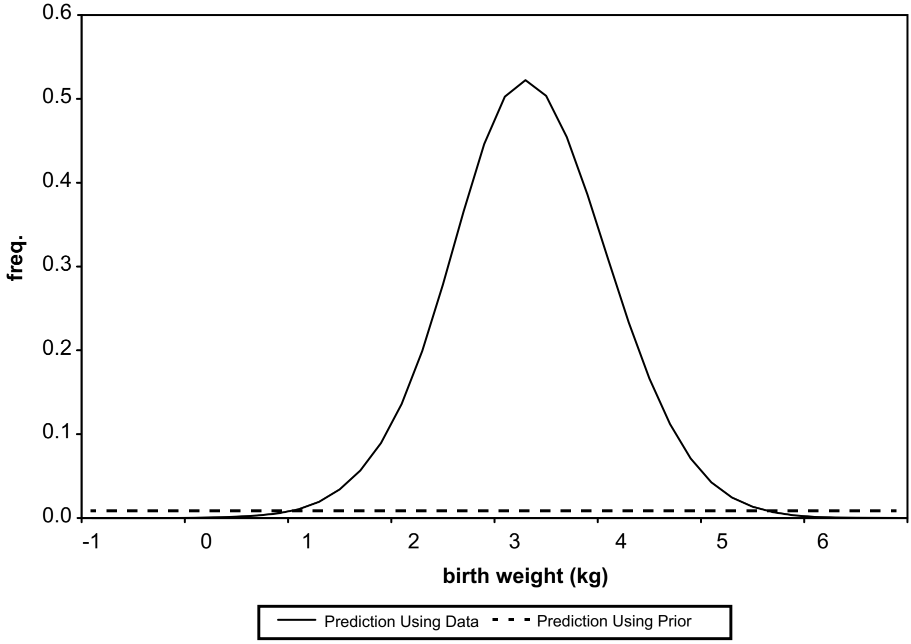

Fig. 1. Predictive density of BW for the reference mother under default prior: $\xi = 1$, $\omega_1 = 0.1936$, $\omega_2 = 9$.

two to obtain $\xi = 2$, $\omega_1 = 0.3872$, $\omega_2 = 18$, respectively. This asymmetrical treatment reflects the fact that the prior predictive distribution from multiplying our default values by four yields an implausible predictive mean for BW.

The Bayes factor continues to overwhelmingly favor the maintained hypothesis H_* over H_A in the two new hyperparameter settings (unreported). And the largest marginal of the data occurs for the tightest prior specification (i.e., $\xi = 0.25$, $\omega_1 = 0.0484$, $\omega_2 = 2.25$). Results under H_A and the tight prior analogous to Tables 6–9 are available from the authors upon request. Not surprisingly, posterior means are of the same sign as before, but have shrunk toward the origin. Posterior standard deviations are reduced as result of the more informative priors.

There are a couple of specific differences worth noting between the results under the default prior and those under the tight prior. Firstly, there is somewhat stronger correlation between three binary input variables (S, D, and PC) under the tight prior than under the default prior. Secondly, there is much stronger positive correlation between ε_S and ε_G under the tight prior than under the default prior. Thirdly, AFQT score (x_{10}), grandmother's education (x_{17}), and not being on time in school at age 14 (x_{18}) look like even stronger instruments under the tight prior than under the default prior.

Table 10 reports the predictive means and standard deviations of the reference mother for all seven endogenous variables corresponding to all three prior hyperparameter specifications and for both our maintained hypothesis H_* and the less restrictive alternative H_A. These predictive moments exhibit fairly little variation over the six specifications.

Table 10
Predictive means (standard deviations) for various specifications

Variable	H_*			H_A		
	$\xi = 0.25$ $\omega_1 = 0.0484$ $\omega_2 = 2.25$	$\xi = 1$ $\omega_1 = 0.1936$ $\omega_2 = 9$	$\xi = 2$ $\omega_1 = 0.3872$ $\omega_2 = 18$	$\xi = 0.25$ $\omega_1 = 0.0484$ $\omega_2 = 2.25$	$\xi = 1$ $\omega_1 = 0.1936$ $\omega_2 = 9$	$\xi = 2$ $\omega_1 = 0.3872$ $\omega_2 = 18$
S	0.5578 (0.4966)	0.5444 (0.4980)	0.5608 (0.4963)	0.5276 (0.4992)	0.5738 (0.4945)	0.5746 (0.4944)
D	0.5052 0.5000	0.5382 (0.4985)	0.5320 (0.4990)	0.5288 (0.4992)	0.5690 (0.4952)	0.5726 (0.4947)
PC	0.5794 (0.4937)	0.6154 (0.4865)	0.6486 (0.4774)	0.5758 (0.4942)	0.6220 (0.4849)	0.6490 (0.4773)
WG	9.698 (8.091)	10.01 (8.260)	10.16 (8.282)	9.779 (7.884)	9.794 (8.242)	9.853 (8.225)
G	38.93 (2.748)	38.81 (2.821)	38.82 (2.839)	38.82 (2.832)	38.76 (2.927)	38.68 (2.964)
BL	50.68 (3.415)	50.79 (3.273)	50.79 (3.289)	50.91 (3.399)	50.89 (3.289)	50.91 (3.287)
BW	3.205 (0.7111)	3.240 (0.7115)	3.233 (0.6966)	3.034 (0.7327)	3.029 (0.7125)	3.026 (0.7228)

The predictive densities of BW for the tight prior, and its corresponding posterior, are shown in Fig. 2.

6. Discussion

The analysis here has demonstrated the feasibility and practicality of our new model of the birth process. Essentially the posterior predictive view of BW suggested by our results is one where the four inputs are determined jointly and dependently among *S*, *D*, and PC, but independently of WG, and then the outputs are determined as follows. *S* has negative systematic correlation with *G*. *D* and PC appear to have no sizeable systematic effect on *G*, BL, or BW. *G* appears dependent on the exogenous size of the

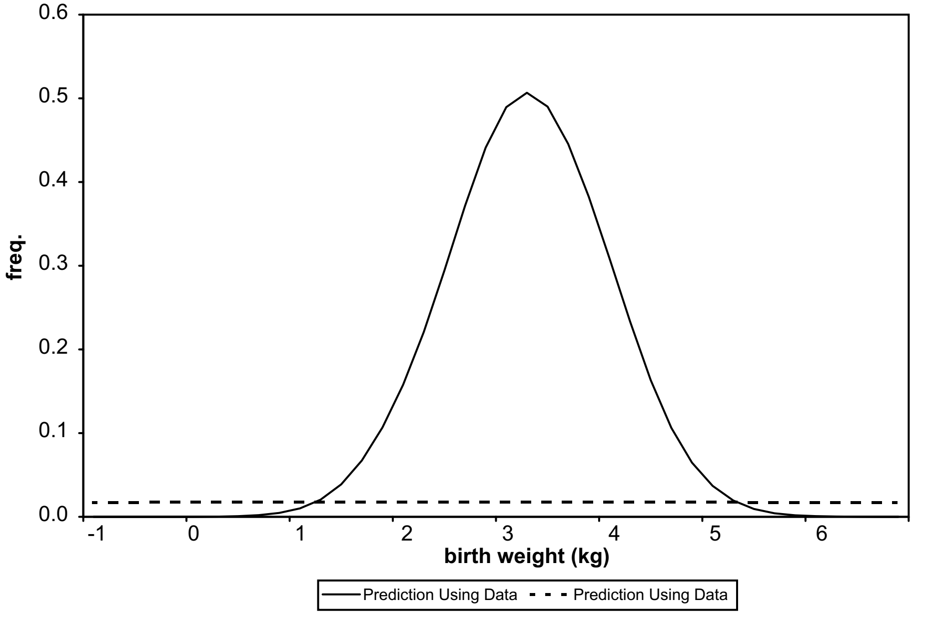

Fig. 2. Predictive density of BW for the reference mother under default prior: $\xi=0.25$, $\omega_1=0.0484$, $\omega_2=2.25$.

mother. Except for the sizeable and positive correlation between the unexplained parts of S and G, there seem to be no unexplained common effects between the birth inputs and outputs. Moreover, G appears dependent on the exogenous size of the mother. BL is affected by the inputs mainly through WG. BW is affected by the inputs through their effects on G. Except for maternal weight, there is little correlation between the remaining exogenous variables and BW.

There are, however, many data considerations that limit our analysis. Firstly, there are issues regarding measurement error in the data we used (e.g., gestational age). Secondly, the small sample size has limited the sharpness of our results. Nonetheless, we feel we have demonstrated well the bottom line of our analysis: the predictive density of BW, (34), as illustrated in Figs. 1 and 2 for the reference mother.

Acknowledgements

The authors gratefully acknowledge the research support of the Social Sciences and Humanities Research Council of Canada, and helpful comments by John Geweke, Gary Koop, Arnold Zellner, three referees, and an associate editor on earlier versions of the paper. We are solely responsible for any errors.

References

Albert, J., Chib, S., 1993. Bayesian analysis of binary and polychotomous response data. Journal of American Statistical Association 88, 669–679.

Carter, R.A.L., Nagar, A.L., 1977. Coefficients of correlation for simultaneous equation systems. Journal of Econometrics 6, 39–50.

Chib, S., 1992. Bayes inference in the Tobit censored regression model. Journal of Econometrics 51, 79–99.

Chib, S., Greenberg, E., 1995. Understanding the Metropolis–Hastings algorithm. American Statistician 49, 327–335.

Chib, S., Greenberg, E., 1998. Analysis of multivariate probit models. Biometrika 85, 347–361.

Chib, S., Jeliazkov, I., 2001. Marginal likelihood from the Metropolis–Hastings output. Journal of the American Statistical Association 96, 270–281.

Cowles, M.K., Carlin, B.P., 1996. Markov chain Monte Carlo convergence diagnostics: a comparative review. Journal of American Statistical Association 91, 883–904.

Geweke, J., 1991. Efficient simulation from the multivariate normal and Student-*t* distributions subject to linear constraints. Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface, pp. 571–578.

Geweke, J., 1999. Using simulation methods for Bayesian econometric models: inference, development, and communication. Econometric Reviews 18, 1–126 (with discussion and rejoinder).

Institute of Medicine, 1985. Preventing Low Birthweight. Division of Heath Promotion and Disease Prevention, Committee to Study the Prevention of Low Birth Weight, National Academy Press, Washington, DC.

Li, K., 1996. Essays in Bayesian financial econometrics. Ph.D. Thesis, University of Toronto.

Li, K., 1998. Bayesian inference in a simultaneous equation model with limited dependent variables. Journal of Econometrics 85, 387–400.

Li, K., Poirier, D.J., 2000. An econometric model of birth inputs and outputs: a detailed report. Working Paper, University of British Columbia, available at http://finance. commerce.ubc.ca/research/abstracts/UBCFIN00-3.html.

Li, K., Poirier, D.J., 2001a. An econometric analysis of the birth process by racial/ethnic groups. In: George, E. (Ed.), Bayesian Methods with Applications to Science, Policy, and Official Statistics: Selected Papers from ISBA 2000. Eurostat, Luxembourg, pp. 321–330.

Li, K., Poirier, D.J., 2001b. Using the National Longitudinal Survey of Youth in the US to study the birth process: a Bayesian approach. Research in Official Statistics 4, 127–150.

Li, K., Poirier, D.J., 2002. Bayesian analysis of an econometric model of birth inputs and outputs. Journal of Population Economics, forthcoming. Available at http://finance.commerce.ubc.ca/∼kaili/BW_JPE.pdf.

McCulloch, R., Rossi, P., 1994. An exact likelihood analysis of the multinomial probit model. Journal of Econometrics 64, 207–240.

Meis, P.J., Goldenberg, R.L., Mercer, B.M., et al., 1998. The preterm prediction study: risk factors for indicated preterm births. American Journal of Obstetric Gynecology 178, 562–567.

Permutt, T., Hebel, J.R., 1989. Simultaneous-equation estimation in a clinical trial of the effect of smoking on birth weight. Biometrics 45, 619–622.

Poirier, D.J., 1995. Intermediate Statistics and Econometrics: A Comparative Approach. MIT Press, Cambridge.

Poirier, D.J., 1998. Modeling birth weight: US stylized facts and a window for viewing them. Working Paper, University of Toronto.

Shiono, P.H., Behrman, R.E., 1995. Low birth weight: analysis and recommendations. The Future of Children 5, No. (1) (http://www.futureofchildren.org/lbw/02lbwana.htm).