**Modelling birth outcomes in New Zealand**

Chris Schilling, Martin Spielauer, and Bill Kaye-Blake

NZIER, Wellington, New Zealand. Corresponding author: chris.schilling@nzier.org.nz.

**Abstract**

A healthy start to life is a critical predictor of later life outcomes. New Zealand ranks poorly in early life outcomes. A new longitudinal study, Growing Up in New Zealand, is compiling an up-to-date dataset of child outcomes and drivers. We use the first two waves of this dataset, which capture the condition and behaviours of the mother and partner during pregnancy, and subsequent births, to develop a microsimulation model of birth outcomes in New Zealand, GUiNZ-Sim. We confirm that GUiNZ-Sim accurately reproduces the survey results. This underpins the validation of a forth-coming population-wide model.

**INTRODUCTION**

*Background*

Policy makers, epidemiologists and scientists have long been aware of the importance of a healthy start to life [Heckman, 2008]. New Zealand performs poorly in many early life metrics, relative to the OECD [OECD, 2009]. Improving the early life outcomes of New Zealand children is a high priority for the government – with ‘A better start’ the second of the government’s 2013 ten National Science Challenges[[1]](#footnote-1).

Growing Up in New Zealand (GUiNZ) is a new longitudinal survey that tracks almost 7,000 New Zealand children and their families [Morton et al, 2010]. The first wave of GUiNZ commenced in 2008 and captured data on the condition and behaviours of the mother and partner during the pregnancy [Morton et al, 2010]. The second wave of data collection captured the birth outcomes, and completed initial explanatory econometric analysis of key drivers [Morton et al, 2012] .

The GUiNZ dataset provides a unique opportunity to develop a policy-relevant microsimulation model of New Zealand children, based on up-to-date New Zealand-specific data. This paper documents the first steps in the development of such a model, GUiNZ-Sim. The aim of GUiNZ-Sim is to turn the wealth of information being collected by the GUiNZ survey into a robust microsimulation model that can be used to analyse various ‘what-if’ policy scenario questions.

*Microsimulation methodology*

Microsimulation is an advanced modelling technique that performs highly detailed analysis at the individual or ‘micro’ level. Rather than using simple population averages, a microsimulation model contains a large number of synthetic ‘people’, called agents. These agents use information, make decisions and experience consequences. They and their offspring may experience negative outcomes as a result of these choices. A microsimulation model uses information from a wide variety of sources to put probabilities on these life events. In the example we present here, we use probabilities derived from the GUiNZ dataset to model how birth outcomes result from parental attributes and behaviours.

The most prominent user of microsimulation in policy analysis is Statistics Canada, who over 20 years have developed a range of health-related microsimulation models. These models include risk factor exposures (such as drinking and smoking), health outcomes (mortality and morbidity) and the demographic characteristics of the Canadian population. Some examples are the Population Health model (POHEM) which focuses on risk attributes and associated chronic diseases such as cancer and asthma; the cancer risk management model which is a specific cancer model that can evaluate various cancer control strategies such as prevention, screening and treatment for lung and colorectal cancer; and the physical activity model which was jointly developed with the Public Health Agency of Canada to investigate how physical activity impacts on life course health outcomes[[2]](#footnote-2).

*Overview of our approach*

Our approach is to develop a cohort model based on the GUiNZ dataset that accurately reproduces the observed birth outcomes. The GUiNZ-Sim cohort model creates a synthetic reproduction of the mothers in the GUiNZ cohort, and the attributes of their pregnancies. The model contains a birth module that probabilistically determines the birth outcomes of each pregnancy, based on the findings of the detailed multivariate analysis of birth outcomes in the GUiNZ dataset.

The development of the GUiNZ cohort model is the first of three planned stages, specifically designed to ensure the model is accurately representing the linkages between pregnancy attributes and birth outcomes. In the second stage, we will develop a population model based on the empirical relationships and calibration of the GUiNZ-Sim cohort model, but applied to the wider New Zealand population. This will deliver a population-wide model of birth outcomes as a function of parental attributes, based on the latest evidence from the GUiNZ dataset.

The model being developed in stage one and two is static: it can be used to estimate the impact of policy interventions on birth outcomes for a given year. In stage three, we plan to make the model dynamic. The benefit of a dynamic microsimulation model is that it can incorporate entire life paths. This is particularly pertinent for early-life analysis, as early-life outcomes are a strong predictor of later life outcomes. The third stage will extend the model in two ways:

1. Extend the model from birth outcomes to early-childhood outcomes as new waves of the GUiNZ survey become available.
2. Extend the model from early-childhood outcomes to later-life outcomes using international research. The International Healthy Start to Life (ISHLP) project has already commenced research in this regard, including collaboration with other longitudinal data surveys, and summary research [for example see Pacheco et al 2013].

Ensuring the GUiNZ-Sim cohort model is documented and accurately reproducing the GUiNZ dataset will assist with validation of the dynamic GUiNZ-Sim population model. We continue now with documentation of the static GUiNZ-Sim cohort model.

**GUiNZ-Sim COHORT MODEL**

*Input data*

The GUiNZ dataset contains 6,822 children and their families. The children were recruited before their birth, so that the pre-natal influence of families and environments could be investigated. The study covers multiple domains that influence child development, including family, societal context, education, health and wellbeing, and culture and identity. It is fully documented in a number of reports including Morton et al (2012) and Morton et al (2010). The key parental and pregnancy attribute data are shown in Table 1.

We note that the GUiNZ dataset is not representative of the wider New Zealand population. It is over-sampled in low income Asian and Pacific groups[[3]](#footnote-3). This is a strength of the dataset rather than a weakness, as Rothman et al (2012) point out: a representative cohort may not have the sample size to identify statistically significant relationships or to narrow confidence intervals for key relationships of interest, typically around our poor and diverse population. By contrast, the GUiNZ dataset is over-sampled in these sub-populations specifically to allow detailed econometric analysis.

*Model*

The model is a standard, continuous-time microsimulation model implemented in the Modgen microsimulation software, licenced from Statistics Canada. The starting population of the model is synthetically generated based on data from the GUiNZ survey. It is disaggregated across sex, ethnicity, age and household income.

As a continuous-time model, the timing of the events in each agent’s life is endogenously calculated within the model. Once the event occurs, the model then calculates the action or outcomes that occur at each event. Typically major life events include at a minimum pregnancy, birth, emigration and mortality. Each major event is given its own sub-module within the modelling code, to improve clarity of coding and assist with debugging.

The pregnancy sub-module calculates the timing of the pregnancy, and assigns each pregnant woman pregnancy attributes. For the GUiNZ-Sim cohort model, all of the female population are already pregnant, so the timing of the pregnancy event becomes redundant. Pregnancy attributes that are assigned include risk factors (smoking, drinking, physical activity and diet), health conditions (hypertension, pre-eclampsia and gestational diabetes) and parity (number of prior pregnancies), all based on the prevalence rates observed in the GUiNZ dataset.

The birth sub-module calculates the timing of the birth event, which is analogous to the gestational length of each pregnancy. This is based on the distribution of gestational age observed in the birth wave of the GUiNZ dataset. Birth attributes are calculated based on the multivariate analysis completed in Morton et al (2012) which explained birth weight from mother and pregnancy attributes (full model description, beta coefficient and standard errors are presented in that paper). The marginal impact on birth weight $∆bw$ of each explanatory attribute *a* for each agent *n* is calculated probabilistically using a random normal number generator *r* with mean of 0 and standard deviation of 1, according to the following formula:

$$∆bw\_{n,a}= β\_{a}+ r\_{n}.SE\_{a}$$

Where β and *SE* are the beta coefficients and standard error results from the multivariate analysis in Morton et al (2012). For example, maternal smoking is shown to have an impact on birth weight of -167.37 grams with a standard error of 28.80 grams. The marginal impact of maternal smoking for a given agent will therefore most likely fall between -138.57 and -196.17 grams, however there is a chance that the impact could be greater or smaller than this. The probabilistic nature of the model captures the variability of impacts across agents.

*Results and discussion*

We find that the GUiNZ-Sim cohort model successfully reproduces birth outcomes between low (<2500 grams), average (2501-4000 grams) and high birth weights (>4001 grams) as shown in Figure 1. This is the primary aim of the model as low birth weight in particular are still shown to be an important predictor of later life outcomes (see for example McGovern, 2013).

At a more granular level, with birth weight disaggregated across 500 gram bins, we begin to see some discrepancies between the GUiNZ-Sim model birth weights and the survey birth weights (Figure 2). In particular, the model under-predicts the number of births in the 3501-4000 gram range but over-predicts the number of births 2501-3000 gram range.

Figure 3 compares average birth weights across ethnicities for the survey versus the model, with the error bar on the survey columns representing plus/minus one standard deviation. We find that the model under-predicts Asian birth weights but over-predicts Pacific birth weights, however all model results are within the confidence limits of the survey data.

Our results provide confidence that the GUiNZ-Sim cohort model is accurately reproducing the GUiNZ survey. This is to be expected given the model inputs and coding. However it is useful first step when considering the verification and validation of the forth-coming dynamic GUiNZ-Sim population model.

**REFERENCES**

Heckman, J.J. 2008. Schools, skills and synapses. IZA Discussion Paper No. 3515. May 2008.

McGovern, Mark E. 2013. Still Unequal at Birth: Birth Weight, Socio-economic Status and Outcomes at Age 9. ***The Economic and Social Review*** *44*, no. 1, Spring, p. 53–84, September.ISSN 0012-9984. Available at: <<http://www.esr.ie/article/view/63>>.

Morton, S.M.B., Atatoa Carr, P.E., Bandara, D.K., Grant, C.C., Ivory,

V.C., Kingi, T.R., Liang, R., Perese, L.M., Peterson, E., Pryor, J.E., Reese, E., Robinson, E.M.,

Schmidt, J.M., and Waldie, K.E. 2010. Growing Up in New Zealand: A longitudinal study of

New Zealand children and their families. Report 1: Before we are born. Auckland: Growing Up

in New Zealand.

Morton, S.M.B., Atatoa Carr, P.E., Grant, C.C., Lee, A.C., Bandara,

D.K., Mohal, J., Kinloch, J.M., Schmidt, J.M., Hedges, M.R., Ivory, V.C., Kingi, T.R., Liang,

R., Perese, L.M., Peterson, E., Pryor, J.E., Reese, E., Robinson, E.M., Waldie, K.E., and Wall,

C.R. 2012. Growing Up in New Zealand: A longitudinal study of New Zealand children and

their families. Report 2: Now we are born. Auckland: Growing Up in New Zealand.

OECD. 2009. Doing Better for Children. New Zealand Country Highlights.

Pacheco, G. Hedges, M., Schilling, C., and Morton, S. 2013. Pre and postnatal drivers of childhood intelligence: evidence from Singapore.

Table 1: Prevalence of pregnancy attributes in the GUiNZ survey data

|  |  |
| --- | --- |
| Pregnancy attribute | Prevalence |
| Gestational diabetes | 3.6% |
| Hypertension | 4.1% |
| Pre-eclampsia | 3.0% |
| Smoking during pregnancy | 11% |
| Passive smoking | 7.2% |
| Parity (first pregnancy) | 42% |
| Medium alcohol consumption (1-3 standard drinks per week) | 3.2% |
| High alcohol consumption (4 or more standard drinks per week) | 0.8% |
| BMI > 30 | 19% |

Figure 1: Birth weight histogram: survey versus model

Figure 2: Disaggregated birth weight histogram: survey versus model

Figure 3: Average birth weights by ethnicity: survey versus model

1. See http://www.msi.govt.nz/update-me/major-projects/national-science-challenges/ [↑](#footnote-ref-1)
2. See http://www.statcan.gc.ca/microsimulation/ [↑](#footnote-ref-2)
3. The reader is directed to Morton et al (2013) for a detailed analysis of the representativeness of the GUiNZ dataset. [↑](#footnote-ref-3)