DISCUSSION PAPER SERIES

Infections, Accidents and Nursing Overtime in a Neonatal Intensive Care Unit: A Bayesian Semiparametric Panel Data Logit Model

Marc Beltempo
Georges Bresson
Jean-Michel Etienne
Guy Lacroix

DISCUSSION PAPER SERIES

Infections, Accidents and Nursing Overtime in a Neonatal Intensive Care Unit: A Bayesian Semiparametric Panel Data Logit Model

Marc Beltempo
McGill University Health Centre

Georges Bresson
Université Paris II

Jean-Michel Etienne
Université Paris-Sud

Guy Lacroix
Université Laval and IZA

MARCH 2020

ABSTRACT

Infections, Accidents and Nursing Overtime in a Neonatal Intensive Care Unit: A Bayesian Semiparametric Panel Data Logit Model

The paper investigates the effects of nursing overtime on nosocomial infections and medical accidents in a neonatal intensive care unit (NICU). The literature lacks clear evidence on this issue and we conjecture that this may be due to empirical and methodological factors. We thus focus on a single NICU, thereby removing much variation in specialty mixes such neonatologists, fellows, residents, nurse practitioners that are observed across units. We model the occurrences of both outcomes using a sample of 3,979 neonates which represents over 84,846 observations (infant/days). We use a semiparametric panel data Logit model with random coefficients. The non-parametric components of the model allow to unearth potentially highly non-linear relationships between the outcomes and various policy-relevant covariates. We use the mean field variational Bayes approximation method to estimate the models. Our results show unequivocally that both health outcomes are affected by nursing overtime. Furthermore, they are both highly sensitive to infant and NICU-related characteristics.

JEL Classification:	I1, J2, C11, C14, C23
Keywords:	neonatal health outcomes, nursing overtime, semi-parametric panel data logit model, mean field variational Bayes, random coefficients

Corresponding author:
Guy Lacroix
Department of Economics
Université Laval
Pavillon J.-A.DeSève
1025, avenue des Sciences humaines
Québec G1V OA6
Canada
E-mail: guy.Lacroix@ecn.ulaval.ca

1. Introduction

A large literature documents the effects of neonatal health on a wide range of adult outcomes such as wages, cognitive skills and human capital accumulation (Black et al., 2007; Oreopoulos et al., 2008; Currie et al., 2010; Figlio et al., 2014; Bharadwaj et al., 2018). Neonatal health is commonly proxied by birth weight or gestational age, as both are highly (and spatially) correlated (Neelon et al., 2014). Indeed, both preterm ($<$37 weeks gestation) and low birthweight infants ($<$ 2500 grams) are more likely to develop neurologic, pulmonary, and gross motor impairments than full-term infants (Behrman and Butler, 2007), and are at higher risk of mortality within one year and up to age 17 (Paneth, 1995; Behrman and Butler, 2007).

Most preterm and low birth weight infants are admitted to a neonatal intensive care unit (NICU) upon birth. Frail newborns are at higher risk of contracting a nosocomial infection (Freeman et al., 1990; Vain et al., 2012) which results in increased morbidity and mortality, prolonged lengths of stay, and increased medical costs (Polin et al., 2012).[1] And because they require more care, are more vulnerable to medical incidents such as erroneous medication administration or feeding and equipment malfunctioning (Beltempo et al., 2017). The onset of nosocomial infections and the occurrence of medical incidents may have adverse health effects that can potentially exacerbate health and socioeconomic problems into adulthood. Understanding the mechanisms that lead to these adverse events may help reduce

---

[1]Neonates are at high risk of acquiring health care-associated infections because of impaired host-defense mechanisms, limited amounts of protective endogenous flora on skin and mucosal surfaces at time of birth, reduced barrier function of their skin, use of invasive procedures and devices, and frequent exposure to broad-spectrum antibiotic agents.

1

the private and societal costs of poor neonatal health as well as contribute to cost containment (Evans and Kim, 2006; Russell et al., 2007; Mistry et al., 2009).

Neonatal intensive care units must contend with ever changing caseloads, patient mix and unplanned admissions (Tucker et al. (1999)). Workforce management is thus challenging and nursing overtime is often used to meet required nurse-to-patient ratios (Berney and Needleman (2005); Beltempo et al. (2016)). The increasing use of overtime hours as a labor management strategy has become an important issue across NICUs in Canada (Canadian Association of Paediatric Health Care Centers, 2013; Fallah et al., 2011) and elsewhere (Griffiths et al. (2014)). This is because nursing overtime has been found to be deleterious to adult patients' health (Bae (2013); Haizhen (2014); Cimiotti et al. (2012); Dorrian et al. (2006); Trinkoff et al. (2011)). Yet, the literature linking nursing overtime and neonatal outcomes, in addition to being relatively scant, is inconclusive (see, *e.g.*, Bae and Favry, 2013; Sherenian et al., 2013). While nurse understaffing *per se* is associated with higher infections rates (Rogowski et al., 2013) and mortality (Watson et al., 2016), mandatory staffing (nurse/patient) has been found to have no impact on health outcomes (Evans and Kim, 2006; Sochalski et al., 2008; Cook et al., 2012).

Understaffing must be viewed in relation to capacity and case-mix. Indeed, it is widely acknowledged that NICUs usually operate at or near capacity, if not beyond. Yet, economists have long questioned whether the availability of supply itself may directly lead to additional utilization (Freedman, 2016). In the words of Roemer (1961), "A built bed is a filled bed", or to paraphrase Carroll (2015), "If you build them, they will come". There is ample evidence that newborns at all birth weights

are increasingly likely to be admitted to a NICU, which raises the possibility of overuse of neonatal intensive care in some not-at-risk or low-risk newborns (Grumbach, 2002; Goodman et al., 2002; Harrison and Goodman, 2015). Likewise, there is excessive regional variation in the proportion of newborns admitted to a NICU that can not be explained by variations in birth weight or gestational age alone.[2] It has been suggested that in some cases C-section delivery could be the sole reason for admitting newborns to NICU/ICU, including for observation with low risk-births (Fallah et al., 2011). Admitting low-risk/low-need infants will artificially decrease the nurse/patient ratio while not necessarily jeopardizing the health status of those more in need of intensive care (Freedman, 2016).

Yet, the lack of clear evidence linking nursing overtime, utilization and patient health may also be due to methodological factors (Bae and Favry, 2013; Weinstein et al., 2008). Indeed, most studies use cross-sectional data and contrast health outcomes stemming from heterogeneous units and/or hospitals.[3] Such analyses are likely to omit important unobserved patient characteristics and unit-specific work arrangements. As for NICUs, given that the mix of neonatologists, fellows, residents, nurse practitioners, *etc.* varies greatly across hospitals, singling out the contributions of nursing overtime and utilization on health outcomes is clearly a difficult task. This difficulty is compounded by the fact that the association between the former and the

---

[2]For instance, in Canada the proportion of newborns admitted to a NICU/ICU between 2006–2009 ranged from 5.3% in the Province of Québec to 24.5% in the Province of New-Brunswick. In addition, the proportion of stays that lasted less than 24 hours varied from 9.8% in the Province of Prince-Edward Island to as much as 39.7% in the Province of Alberta (Fallah et al., 2011).

[3]Yet, see Mújica-Mota et al. (2020) for a recent analysis which accounts for heterogeneous causal effects of neonatal care on mortality.

latter is perhaps not a linear cause-effect relationship (Hugonnet et al., 2006).

In this paper, we focus on the CHU de Québec NICU, a tertiary/quaternary referral center with a 51-bed capacity that tends to a population of 1.7 million over a territory of 452,600 km^2 surrounding Québec City, Canada. Focusing on a single unit removes some of the aforementioned variations in specialty mix across NICUs. We study the daily occurrence of health care associated infections and medical incidents/accidents (henceforth HCAI and MA, respectively) among all neonates admitted to the NICU between April 2008 and March 2013. Daily exposure to overtime and regular hours of work, as well as numerous individual and NICU-specific covariates are used to model the onset of the latter two outcomes. We also exploit an important change in workforce arrangement that was implemented in June 2012 and which aimed at reducing overtime hours. Management thus hired 15 full-time registered nurses and converted 10% of existing positions from 8-hour to 12-hour shifts which were exempted of additional overtime hours (Beltempo et al., 2016). We use a flexible semiparametric logit model with random coefficients to quantify the links between the main variables of interest and the two outcomes. The non-parametric components of the model allow to unearth potentially highly non-linear relationships in overtime, regular hours of work and birth weight, and is well-suited to measure the sensitivity of the outcomes to the new workforce arrangement. Given the size and the length of our (unbalanced) sample, we use the recent streamlined mean field variational Bayes estimator proposed by Lee and Wand (2016) which allows fast and efficient estimation of the model parameters.

Section 2 presents the data. Section 3 gives some insights about the mean field

4

variational Bayes estimation method of the semiparametric Logit model with random coefficients. Section 4 gives the results while section 5 concludes.

2. Data and Institutional Arrangement

2.1. The NICU

The CHU de Quebec NICU is a Level-III referral center with a 51 bed capacity. Nurse staffing is determined before each shift according to patient acuity, planned admissions, and elective procedures/tests. When nurses are deemed in shortage, management initially turns to available off-duty nurses. Next, a pool of floating nurses is relied upon. Finally, it resorts to voluntary and mandatory overtime if necessary. Overtime is defined as all hours worked beyond the regular work schedule.[4]

Daily administrative data on overtime and regular hours of work, daily patient census and number of admissions were collected using the local administrative database *Logibec*. This latter is used to manage work shifts and pay schedules. The information on regular hours and overtime is thus quite precise. Information on HCAI was drawn from the infectious disease database *TDR* while information on MA was retrieved from the *Gesrisk* database.[5]

Figure 1 below exhibits the smoothed daily variations in occupancy rates and overtime hours over our entire sample period (April 2008–March 2013). Occupancy rates are expressed in percentage relative to capacity (51 beds). They vary between 72% (35 filled beds) and 113% (58 filled beds). The NICU operates above capacity

---

[4]Overtime occurs whenever a nurse either starts her shift earlier than planned or finishes later than scheduled. Working beyond 16 consecutive hours per day is forbidden.

[5]Reporting the information on the timing as well as the type of MA is mandatory.

41% of the time. Yet this occurs more frequently after the change in the overtime regime implemented by management in June 2012 (vertical dashed line). Indeed, prior to the implementation of the new policy, the NICU operated above capacity 35% of the time. The proportion increased to 71% in the aftermath. This is clearly depicted in the figure.

Overtime hours follow the opposite path. Prior to the implementation of the new regime, average daily overtime hours amounted to 23.6. In the months that followed, it decreased to 18.6. The implementation of the policy occurred at a time when occupancy was relative high and overtime hours relatively low. Recall from

Figure 1: Smoothed Daily Occupation Rates and Overtime hours
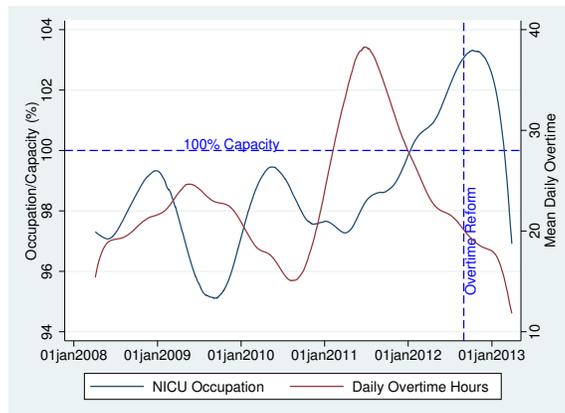
our discussion above that some have noted that infants at all birth weights are increasingly likely to be admitted to a NICU (Harrison and Goodman, 2015). If this it the case, then the health status at admission should be inversely related to the occupancy rates. Indeed, as available beds become fewer, management will naturally prioritize high-risk infants. Figure 2 below investigates this issue. The figure reports the (smoothed) average weight and weight/gestational age at admission by occupancy

6

rate. It also distinguishes between pre and post reform periods.[6] According to

Figure 2: Weight and Weight/Gestational Age at Admission, by Occupancy Rate
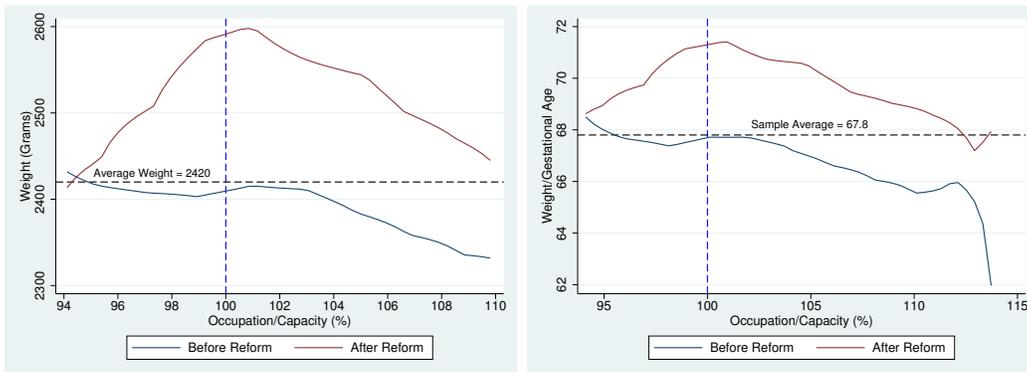
the figure, prior to the reform both the weight and weight/gestational age were relatively independent of the occupancy rate up until a rate of 102%-103%. Above this rate, infants admitted to the NICU had slightly poorer health. Indeed, the average weight is roughly 100 grams lower and the weight/age ratio less by two units between occupancy rates of 102% and 110%.[7] On the other hand, infants admitted in the post-reform period have both a greater average birth weight and Weight/Age ratio. The mean weight difference is 132 grams and the Weight/Age more than 3 units. Both differences are highly statistically different. As in the pre-reform period, the health status exhibits an inverse relation with the occupancy rate. Above 102%, the birth weight and the Weight/Age ratio decrease at the same rate as those admitted prior to the reform. The negative relation between health status and occupancy lends credence to the claim that NICUs may have an incentive to operate

---

[6]The figures depict local polynomial regressions for occupancy rates above 94%. The NICU operates at or above this rate 80% of the time. There are too few observations at lower rates to make valid statistical inference.

[7]3,322 infants were admitted prior to the reform, and only 657 in the aftermath.

at capacity and thus admit low-risk/low-need infants to that end (Freedman, 2016). In the particular case of the Quebec NICU, this may be related to the fact that it operated on average at or below capacity in the pre-reform period and somewhat above capacity in the post-reform period as shown in Figure 1.[8]

2.2. Patient Characteristics and Health Outcomes

Patient characteristics are drawn from the hospital clinical database. The latter includes information on gestational age, birth weight, sex, Apgar score, multiple pregnancies, type of delivery, *etc.* All newborns admitted during the study period were included conditional on having spent at least four days in the NICU. If an infant had more than one episode of bacteremia, these were considered separate events if they occurred more than 14 days apart. The date of the infection was determined as that at which the blood culture was obtained.

The total number of infants admitted in the NICU over our sample period is 7,383. Of those, 3,404 were omitted since their stay was shorter than four days. The motivation for excluding short spells is twofold. First, HCAIs occur whenever a pathogenic organism can be isolated in blood or a cerebrospinal fluid culture. Such cultures usually require at least two to three days to be conclusive. Second, the exact timeline of events that leads to HCAIs is still not well defined in the literature (Polin et al., 2012). Studies focusing on adult and pediatric units suggest that low nurse-to-patient ratios and extensive overtime over a 3-day period may trigger the onset

---

[8]The unit operated above capacity 35.5% of days in the pre-reform period ($\overline{\text{Capacity}} = 98.4$) and more than 75.3% in the post-reform period ($\overline{\text{Capacity}} = 103.4$.) The difference is highly statistically significant.

8

of HCAIs (Cimiotti et al., 2006; Beltempo et al., 2017). In line with this literature, overtime hours on any given day was thus defined as a moving average computed over the preceding 3-day period when the analysis focuses on HCAIs. Such a moving average can only be computed for spells lasting at least 4 days. When analyzing medical accidents, on the other hand, we use daily exposure to overtime hours.

The final sample consists of 3,979 neonates which represents over 84,846 infant/days over the sample period. Table 1 provides descriptive statistics of the main variables used in the model. The first column focuses on infants who did not contract a HCAI or suffered an MA. The next two columns focus on the subsample of infants who suffered an MA and contracted a HCAI, respectively. Not surprisingly, the table highlights the link between poor health and adverse health events. Indeed, infants who suffered a medical accident or contracted an infection had both a have lower gestational age and birth weight than otherwise. They also had a lower Apgar score and were more likely to have been delivered by C-Section. The next two lines focus on the Diagnostic-Related Group (DRG). DRGs categorize patients with respect to the main diagnosis at admission. It takes one of four potential values: 1=mild, 2=moderate, 3=severe, 4=extreme. The next line reports the proportion of infants who were admitted to the NICU following surgery. As previously, infants with an MA or a HCAI were more frail at entry, were much more likely to have undergone surgery and somewhat more likely to be a first delivery. Poor health translates into lengthy hospitalization spells, and much more so for those with an MA or a HCAI. The last line reports the proportion of infants whose spell occurred during of after the implementation of the overtime reform.

9

The next panel of the table focuses on the average daily characteristics of the NICU by outcome subsamples. The only two noteworthy features concerns the hours of work. Indeed, infants who contracted a HCAI or suffered an MA were exposed to slightly more overtime and regular hours of work than otherwise. This is true despite the fact that there were on average no more admissions into the unit, nor was the occupancy rate greater than usual.

The last panel of the table provides a detailed account of the two outcome variables. Overall, 3,513 infants in our sample had an MA or HCAI-free stay in the unit. On the other hand, 300 infants (7.54%) suffered 389 MA events, and 240 of them (6.03%) contracted 272 HCAIs. From the NICU's point of view, the probability of observing an MA or a HCAI in any given day was 21.35% and 14.92%, respectively. This translates into 0.458% and 0.321% when computed daily and per neonate. The above discussion highlights the fact that the link between nursing overtime, case-mix, and capacity is a complex one. The tables and figures provide at best weak *prima facie* evidence to the effect that the adverse health events may be loosely related to hours of work. Yet the influence of other variables needs to be netted out in order to determine the precise link between work schedules and health outcomes, if any. It is highly unlikely that standard econometric methods are capable of unearthing what we suspect are heterogeneous and non-linear conditional means responses. In what follows, we briefly sketch the recent flexible semiparametric logit model with random coefficients proposed by Lee and Wand (2016). The non-parametric components of the model are particularly well-suited to investigate potentially highly non-linear relationships between our two outcomes and overtime and regular hours of work.

3. The semiparametric Logit model with random coefficients

In what follows, we briefly sketch the semiparametric random effects logit model.[9] Consider the probability that the binary outcome (MA or HCAI), y_{it}, for infant i occurs during its t-th day in the NICU. The unbalanced sample is composed of N infants ($i = 1, ..., N$) each observed during T_i days ($t = 1, ..., T_i$). Let y be the ($NT \times 1$) vector of the y_{it} probabilities with $T = \sum_{i=1}^{N} T_i$. Consider the following mixed effects logistic model:

$$y \mid \beta, u \sim \text{Bernoulli}\left(\text{logit}^{-1}\left(X^R \beta^R + Z^R u^R + f\left(X^G\right)\right)\right), \tag{1}$$

where $y \sim \text{Bernoulli}(p)$ is shorthand for the elements of y having independent Bernoulli distributions with parameters corresponding to those in $p(\cdot)$, and $\text{logit}^{-1}(x)$ is shorthand for the logistic distribution, *i.e.* $e^x/(1 + e^x)$. We borrow the approach and notation of Zao et al. (2006) and Lee and Wand (2016). Let X^R be a ($NT \times q^R$) matrix of covariates and Z^R a ($NT \times Nq^R$) block-diagonal matrix of the X_i^R submatrices. X and Z are called the fixed and random effects design matrices associated with β and u.[10] $X_{it,1}^R$ is the intercept and $X_{it,j}^R$, $2 \leq j \leq q^R$ are the other control covariates. The random intercept is defined by the sum ($\beta_1^R + u_{i,1}^R$), the random slope for variable $X_{i,2}$ is the sum ($\beta_2^R + u_{i,2}^R$), *etc.*

---

[9]The Mean Field Variational Bayes estimator is detailed in a companion web appendix.

[10]This terminology is different from that of the classical panel data literature which refers to "fixed" and "random" effects.

The semiparametric additive function is given by

$$f\left(X^G\right) = X^G\beta^G + Z^G u^G = \sum_{l=1}^{L} X_l^G \beta_l^G + \sum_{l=1}^{L} Z_l^G u_l^G. \tag{2}$$

The matrices X^G and Z^G are associated with the fixed effects vectors β^G and u^G. The $(NT \times L)$ matrix X^G contains L covariates that are not already included in X^R. In our specific case, the three variables in X^G are regular and overtime hours of work, and birth weight.[11] The $(NT \times q^L)$ matrix Z^G matrix, with $q^L = \sum_l^L q_l^G$, contains spline basis functions of the three covariates using q_l^G knots, and u^G are the corresponding spline coefficient vectors. Following Lee and Wand (2016), we use transformed cubic O'Sullivan splines (see also Wand and Ormerod (2008)). Therefore,

$$X^R = vec\left(X_1^R, ..., X_N^R\right), \ Z^R = \operatorname*{blockdiag}_{(1 \le i \le N)}\left(X_i^R\right), \beta = \left(\beta^{R'}, \beta^{G'}\right)' \tag{3}$$

$$u = \left(u^{R'}, u^{G'}\right)', X = \left(X^R, X^G\right) \text{ and } Z = \left(Z^R, Z^G\right)$$

$$Cov\left(u^R\right) = I_N \otimes \Sigma^R, Cov\left(u^G\right) = \operatorname*{blockdiag}_{(1 \le l \le L)}\left(\sigma_{u_l}^2 I_{q_l^G}\right)$$

$$\text{and } \Psi = \begin{pmatrix} Cov\left(u^R\right) & 0 \\ 0 & Cov\left(u^G\right) \end{pmatrix} = \begin{pmatrix} I_N \otimes \Sigma^R & 0 \\ 0 & \operatorname*{blockdiag}_{(1 \le l \le L)}\left(\sigma_{u_l}^2 I_{q_l^G}\right) \end{pmatrix}.$$

Σ^R is an unstructured $\left(q^R \times q^R\right)$ covariance matrix, $\otimes$ denotes the Kronecker product

---

[11]This specification is based on the results of a generalized mixed effects regression tree model (GMERT) using the same database as in Beltampo et al. (2019).

and I_N is an $(N \times N)$ identity matrix. $\sigma_{u_l}^2$ is the penalized parameter for the spline coefficients $\left\{ u_l^G, 1 \leq l \leq L \right\}$ and Ψ is the random effects covariance matrix associated with the vectors u^R and u^G.

This semiparametric panel data model with random coefficients is a complex specification and the methods for estimating such a specification are few. Although the nonlinear panel data literature provides examples of either Logit models with random coefficients (for instance Moon et al. (2017)) or semi- (non-) parametric Logit models (see Lewbel (2000), Honoré and Lewbel (2002) or Ruppert et al. (2003)), examples of models that encompass both features are scarce. Yet, including semiparametric components into a panel data Logit model with random coefficients yields a very rich model that should prove useful in many circumstances in economics. Unfortunately, standard MCMC Bayesian techniques such as Gibbs sampling are inadequate with such models since they are typically computationally prohibitive, may suffer from poor mixing and may not scale well when applied to models that require the inversion of large sparse covariance matrices (as in our case). Fortunately, Lee (2016) and Lee and Wand (2016) have proposed to use a variational Bayesian method known as "mean field variational Bayes approximation" (henceforth MFVB) to estimate such models. MFVB consists in a set of tools which provide a good approximation of the posterior distributions of the parameters. Because the posterior distribution is approximated, MFVB is much faster than traditional Bayesian methods and can afford to tackle large models such as ours.[12]

---

[12]See Appendix A for the technical details.

4. Estimation Results

Our estimation strategy consists in specifying first the most general version of the model and then testing whether a more parsimonious version performs just as well. We thus begin with the semiparametric Logit model with random coefficients $(Z^G = \text{blockdiag}\left(X_i^R\right)$ with $1 \leq i \leq N)$. The model has $Nq^R(= 3,979 \times q^R)$ random coefficients $(\beta_j^R + u_{i,j}^R)$ with $j = 1, ..., q^R$ and $i = 1, ..., N$, which is huge and time consuming. Using a Wald test, we next test the null hypothesis of constant slopes and random intercept, *i.e.* $H_0 : u_{i,j}^R = 0$ for $j = 2, ..., q^R$.[13] The Wald test follows a $\chi^2_{N(q^R-1)}$ and the null hypothesis could not be rejected in either HCAI and MA models. Our analysis thus focuses on the restricted semi-parametric Logit model with a random (intercept) effects, $(\beta_1^R + u_{i,1}^R)$, and constant slopes (β_j^R) for the X^R covariates.

4.1. Medical Accidents

The parameter estimates of the probability that a medical accident occurs are reported in Table 2. The table is divided into three sections. The first reports the parameters β^R, the second one focuses on β^G, and the third reports the panel-level variance component, $\sigma^2_{u^R}$, and the implicit proportion the total variance due to the panel-level variance component, $\rho_{\sigma^2_{u^R}}$. The table also reports the associated odds ratios as well as their 95% confidence intervals.[14]

---

[13]As shown in the online supplement, the Wald test in the context of the MFVB estimation of such semiparametric models has optimal empirical size.

[14]Details on the estimation of the parameter estimates, their standard deviations, and the algorithms used are presented in the Appendix A. Note that the MFVB approximates the posterior distribution by the product of the so-called optimal q-densities which minimize the Kullback-Leibler

According to the table, most parameter estimates are different from zero at the 5% level and most bear very small standard deviations. In addition, the signs of most coefficients are consistent with the previous literature. We first focus on the characteristics of the infant and the delivery. Infants who contracted a nosocomial infection during their spell in the NICU (`Prior Infection`) are much less likely to be victim of a medical accident in the aftermath. This is perhaps because they have become more vulnerable and thus require more attentive care. `Gestational Age` is positively related to the occurrence of a MA, which is perhaps surprising. Yet, since we are controlling for birth weight, a longer gestational age may in fact be associated with a poorer health status. In other words, lower weight-for-age ratio infants may require more care or handling which may result in an increased likelihood of MA. Whether the infant was admitted in the NICU at birth or transferred from another unit or hospital (`Birth vs Transfert`) has no impact on the occurrence of a MA. The same applies to the `Sex` of the infant. Infants who were delivered through `C-Section`, `Twins` and those with a high `Apgar Score` are also less likely to experience a MA. Finally, those who were admitted to the NICU following a `Surgical` intervention and those whose status was deemed highly severe conditional on their `DRG` code, are much more likely to be victim of a MA.

The previous parameter estimates relate to factors that are infant-specific. The

---

divergence. This minimization of the Kullback-Leibler divergence is thus equivalent to maximizing the "Evidence Lower Bound (ELBO)". Convexity properties of the MFVB algorithm guarantees quick convergence to at least a local optimum. In our case, convergence was easily reached within 500 cycles. Computation required 5791.23 seconds of computing time, which is orders of magnitude faster than MCMC. The estimations were conducted using R version 3.3.2 on a 24-worker machine in the department of economics at Université Laval. Some elements of the R code are available in the supplementary material of Lee and Wand (2016).

next set of estimates focus on two factors that are unit-specific. These measure the impact of daily variations in `Unit Occupancy` and `Daily Admissions` on the probability of observing a MA for a particular infant. It is implicitly assumed that all infants are equally exposed to the same "dose" of the latter two. According to the parameter estimates, only `Occupancy` has a statistically significant, albeit very small and negative, effect. This result is consistent with those of Beltempo et al. (2017) using similar data but a different estimation strategy, and for which unit occupancy was found to have no impact on health outcomes.

Columns 3–6 of the table report the odds-ratios of the X^R variables, their standard deviation as well as their 95% credible intervals. These are computed as follows. Assume the X_j^R covariate varies by an amount δ, while the others remain constant. Then, the odds ratio and its standard deviation are given by $OR\left(X_j^R\right)_\delta = \exp\left(\delta\hat{\beta}_j\right)$, and $\sigma_{OR\left(X_j^R\right)_\delta} = \hat{\sigma}_{\beta_j} \exp\left(\delta\hat{\beta}_j\right)$. For dummy variables, we consider $\delta = b-a$ with $a=0$ and $b=1$. The odds ratio is given by the ratio of the odds of an event occurring with the risk factor $(X_j^R = b)$ to the odds of it occurring with the risk factor $(X_j^R = a)$. When the event $y_{it} = 1$ is rare as in our case, then $OR\left(X_j^R\right)$ is approximately equivalent to the relative risk $RR\left(X_j^R\right)$.[15] According to Table 2, and using $\delta = 1$, the lowest odds ratios concern `DRG severity` (OR $\in$ [0.33-0.68]), `Prior Infection` (OR = 0.59) and `Apgar Score` (OR = .76), while the highest odds ratios relate to `Gestational Age` (OR = 1.06), `Overtime Reform` (OR = 1.29) and `Surgical vs Medical` (OR = 1.41). Thus the probability that a MA occurs on

---

[15]The relative risk is the ratio of the probability of the outcome with the risk factor $(X_j^R = b)$ to the probability of the outcome with the risk factor $(X_j^R = a)$.

any given day and for any given infant varies greatly with the conditioning variables at both the infant and the NICU levels. For instance, infants who contracted a `Prior Infection` in the NICU are 41% less likely to experience a MA. On the other hand, infants admitted after the `Overtime Reform` are 29% more likely to do so.

Recall that the `Overtime Reform` that was implemented in June 2012 consisted in hiring 15 full-time registered nurses and in converting 10% of existing positions from 8-hour to 12-hour shifts that were exempted from overtime hours. As outlined above, this has resulted in additional medical accidents, *ceteris paribus*. While the reform has resulted in fewer daily overtime hours (from 24.8 to 21.01, see Figure 1), it has also resulted in more average daily regular hours of work (from 514 to as many as 554 in the post-reform period). The new mix of overtime and regular hours of work may have have resulted in fewer MAs overall. What the parameter estimate shows is that, *ceteris paribus*, the `Overtime Reform` in itself has had a positive impact on the probability of observing a MA.

We further investigate the links between hours of work and MA in the next panel of the table which reports the fixed parameters β^G of the semiparametric additive function. According to these, `Daily Overtime` and `Daily Regular` hours of work have a positive and statistically significant effect on MA. In addition, as expected, a greater `Birth Weight` translates into a lower probability of a MA. The main benefit of using a semi-parametric model is its ability to compute varying and potentially non-linear odds-ratios for each variable in X^G. Unfortunately, their computation as well as that of their standard deviations are quite involved and so

must be bootstrapped (see Appendix B).[16] Figure 3 draws two-dimensional contour plots for different combination of X^G variables, conditional on a given value of a third variable. Thus, Figure 3a shows the sensitivity of the odds-ratio of the probability of a MA as `Daily Regular Hours` and `Daily Overtime Hours` vary, conditional on the 25th decile of `Birth Weight` (1080 grams). For such low weight infants, the odds ratios increase very rapidly from 1 to 4.5. The odds ratios are very sensitive to the `Daily Overtime Hours` for values above 60. Figure 3b plots the odds ratios of the probability of MA as `Birth Weight` and `Daily Overtime Hours` vary, conditional on the 75th of `Daily Regular Hours` (558 hours). The contour surfaces are non-linear and the odds-ratios peak at values of `Birth Weight` below 1,000 grams. The probability of observing a MA for infants of mean weight values is much less sensitive to *Daily Overtime Hours*. Finally, Figure 3c reports the odds-ratios between `Daily Regular Hours` and `Birth Weight`, conditional on the 75th value of `Daily Overtime Hours` (34 hours). Once again, the odds-ratios are non-linear. The probability of observing a MA increases rapidly with regular hours of work but this is mostly true for low birth weight infants.

The parameter estimates show that the occurrence of a MA is sensitive to many individual and NICU-specific factors. The semi-parametric function underlines the fact that an infant is more likely to experience a MA the lower his birth weight, and whenever daily regular or overtime hours are well above average.

---

[16]Figures A1–A12 of the online supplement reports the conditional contours plots of the odds ratios of the probability of MA conditional of the couple (`Daily Regular Hours`, `Daily Overtime Hours`) for 3 deciles of `Birth Weight` (10, 50, 90), of the couple (`Daily Regular Hours`, `Birth Weight`) for 3 deciles of `Daily Overtime Hours` (10, 50, 90), and of the couple (`Daily Overtime Hours`, `Birth Weight`) for 3 deciles of `Daily Regular Hours` (10, 50, 90).

4.2. Health Care Associated Infections

Estimation results for HCAI are reported in Table 3 whose setup is identical to the previous one.[17] HCAIs have been much more investigated in the literature than MA presumably because nosocomial infections are likely to have long term health and socio-economic consequences (Bharadwaj et al., 2018). As stated earlier, netting out the impact of overtimes hours on HCAIs is a difficult task. This is perhaps why the literature is inconclusive on this issue (Bae and Favry, 2013; Weinstein et al., 2008).

According to the table, a number of infant and birth-specific factors have a statistically significant effect on the probability of observing a HCAI. Thus, having experienced a MA earlier in the spell (`Prior MA`) decreases the probability considerably. This may be linked to the fact that greater caution is exercised when caring for these infants. The associated odds-ratio is equal to 0.63, so that infants who experienced a MA during their stay in the NICU are 37% less likely to contract a nosocomial infection in the aftermath. Likewise, being admitted to the NICU at birth rather than being transferred from the nursery (`Birth vs Transfer`) may be a proxy for frailty as the probability of contracting a nosocomial infection is much greater in the former case. This result is consistent with the fact that some not-at-risk or low-risk newborns may be transferred to the NICU when occupancy is below capacity (Grumbach, 2002; Harrison and Goodman, 2015). Indeed, according to the odds-ratio infants who are admitted in the NICU at birth are 32% more likely to

---

[17]As with MA, convergence easily obtains within 500 cycles, or 5953.49 seconds of computing time.

contract a nosocomial infection than those who are transferred from another hospital unit. `C-Section` births are also associated with a much lower probability of infection. Recall that many have suggested that `C-Section` delivery could be the sole reason for admitting otherwise healthy infants to the NICU (Fallah et al., 2011). Our result is consistent with this conjecture, and with an odds-ratio of 0.82, the effect is sizeable. The reason for admission (`Surgical` *vs* `Medical,` `OR=1.36`) and the `DRG-Severity` at admission ($OR \in [0.09, 0.84]$), not surprisingly, are very strong predictors of HCAIs. Indeed, infants who are admitted for surgical reasons are on average 36% more likely to contract a nosocomial infection that those admitted for medical reasons. Likewise, conditional on the `DRG` code, infants whose severity is deemed low are 91% less likely to do so than those whose severity is considered very high. Medium and high severity `DRG` codes also lead to much lower probabilities of contracting a nosocomial infection than `DRG` codes classified as very high.

The only NICU-specific factor which impacts the probability of observing a HCAI is `Daily Admissions (OR=1.04)`, although the effect is relatively small. This is in line with the findings of Beltempo et al. (2017) using other data and a simple logistic regression and which found `Daily Admissions` to have little or no impact on the occurrence of a HCAI. Interestingly, the `Overtime Reform` has had no impact on the probability of observing a HCAI, unlike what was found above for MAs. This is not to say that overtime hours have no impact on HCAIs, quite to the contrary. Indeed, the next panel of the table reports the fixed parameters β^G of the semi-parametric additive function. Interestingly, `Daily Overtime Hours` has a significant and sizable positive impact on the occurrence of HCAIs, whereas `Daily Overtime Hours` does

not. `Birth Weight`, as expected, has the opposite effect.

We next investigate the sensitivity of the probability of observing a HCAI with respect to the three components of the semi-parametric function.[18] Figure 4 draws conditional two-dimensional contour plots for the same combinations of X^G variables as in Figure 3. Once again, the figure underlines the benefits of using a flexible semi-parametric specification. Indeed, Figure 4a plots the contours of the odds-ratios between `Daily Regular Hours` and `Daily Overtime Hours`, conditonal on the 25th decile of `Birth Weigh`t. As with MA, the patterns are relative linear and increase steadily as `Daily Overtime Hours` increase. Note that the odds-ratios are little sensitive to the `Daily Regular Hours`, consistent with the fact its fixed parameter estimate, β_{RH}^G, is not statistically significant. On the other hand, the odds-ratios are very sensitive the increases in `Daily Overtime Hours`, and mostly so above 60 hours. Indeed, increasing the number of `Daily Overtime Hours` from 60 to 120 translates into a threefold increase in the odds-ratios. Figure 4b plots the contour of the odds-ratios between `Birth Weight` and `Daily Overtime Hours`, conditional on the 75th of `Daily Regular Hours` (558 hours). The contour surfaces are non-linear and the odds-ratios peak at values of `Birth Weight` below 1,000 grams. The odds-ratios for infants of weight above the mean sample of 2,400 grams are much less sensitive to overtime hours. Indeed, increasing the `Daily Overtime Hours` from 60 to 120, increases their odd-ratios from 1 to 1.5. Finally, Figure 4c reports the odds-ratios

---

[18]Figures A13–A26 of the online supplement provides conditional contour plots of the odds ratios of the probability of HCAI for the couple (`Daily Regular Hours`, `Daily Overtime Hours`) for 3 deciles of `Birth Weight` (10, 50, 90), for the couple (`Daily Regular Hours`, `Birth Weight`) for 3 deciles of `Daily Overtime Hours` (10, 50, 90) and for the couple (`Daily Overtime Hours`, `Birth Weight`) for 3 deciles of `Daily Regular Hours` (10, 50, 90).

between `Daily Regular Hours` and `Birth Weight`, conditional on the 75th value of `Daily Overtime Hours` (34 hours). Because the parameter estimate associated with `Regular Hours`, β_{RH}^{G}, is not statistically significant, the contour surfaces of the odds ratios increases linearly by strata. For infants at or below average weight ($\leq$ 2,400 grams), the risk of HCAI increases considerably as birth weight decreases, regardless of regular workload.

5. Conclusion

Most preterm and low birth weight infants are admitted to a neonatal intensive care unit (NICU) upon birth. Frail newborns are at higher risk of contracting a nosocomial infection (Freeman et al., 1990; Vain et al., 2012) which may result in increased morbidity and mortality, prolonged lengths of stay, and increased medical costs (Polin et al., 2012). And because they require more care, are more vulnerable to medical incidents such as erroneous medication administration or feeding and equipment malfunctioning (Beltempo et al., 2017).

NICUs are complex entities that are challenging from a managerial point of view. Unplanned admissions, random patient mixes, ever changing caseloads, *etc.* require a particularly flexible workforce. In order to meet to meet required nurse-to-patient ratios (Berney and Needleman, 2005; Beltempo et al., 2016), management often turns to nursing overtime (Canadian Association of Paediatric Health Care Centers, 2013; Fallah et al., 2011; Griffiths et al., 2014). Reliance on nursing overtime has become an important policy debate because it has been found to be deleterious to adult patients' health (Bae (2013); Haizhen (2014); Cimiotti et al. (2012); Dorrian

22

et al. (2006); Trinkoff et al. (2011)). Yet, the literature linking nursing overtime and neonatal outcomes, in addition to being relatively scant, is inconclusive (see, *e.g.*, Bae and Favry, 2013; Sherenian et al., 2013).

We conjecture that the lack of clear evidence linking nursing overtime, utilization and patient health may also be due to methodological factors. Indeed, most studies use cross-sectional data and contrast health outcomes stemming from heterogeneous units and/or hospitals. Such analyses are likely to omit important unobserved patient characteristics and unit-specific work arrangements. As for NICUs, given that the mix of neonatologists, fellows, residents, nurse practitioners, *etc.* varies greatly across hospitals, singling out the contributions of nursing overtime on health outcomes is clearly an empirically difficult task. This difficulty is compounded by the fact that the association between the latter two is perhaps not a linear cause-effect relationship (Hugonnet et al., 2006).

In this paper, we focus on a single tertiary NICU with a 51-bed capacity. Focusing on a single unit removes some of the aforementioned variations in specialty mix across NICUs. We study the daily occurrence of health care associated infections and medical incidents/accidents among all neonates admitted to the NICU between April 2008 and March 2013. Daily exposure to overtime and regular hours of work, as well as numerous individual and NICU-specific covariates are used to model the onset of the latter two outcomes.

We use a flexible semiparametric logit model with random coefficients to quantify the links between the main variables of interest and the two outcomes. The non-parametric components allow to compute highly flexible odds-ratios between them.

Furthermore, given the size and the length of our (unbalanced) sample, we use the recent streamlined mean field variational Bayes estimator proposed by Lee and Wand (2016) which allows for fast and efficient estimation of the model parameters. Our results provide clear evidence that the onset of nosocomial infections and the occurrence of medical accidents are intimately related to infant-specific and NICU-specific characteristics. In particular, both outcomes are shown to be highly sensitive to nursing overtime. Importantly, the sensitivity of the two outcomes is shown to vary greatly with the mix of regular and overtime hours, as well as with the infant's birth weight. Thus infants at, or above, average birth weight and who are exposed to little overtime are at low risk of contracting a nosocomial infection. On the other hand, low birth infants who are exposed to numerous overtime hours are considerably more at risk. It is thus conceivable that the inconclusiveness of the literature may partly due to the fact that standard regression models focus on mean point estimates. Allowing more flexibility in the model is perhaps better suited to unearth subtle non-linear relationships between outcomes and important policy variables.

Table 1: Sample Means and Standard Deviations

Variable	Accident Infection/ No	Accident Yes	Infection Yes
	NEONATES AT ADMISSION		
Sex (Female=1)(%)	44.37	47.33	43.75
	(0.49)	(0.50)	(0.50)
Gestational Age	35.27	32.00	29.72
	(3.49)	(5.02)	(4.25)
Weight (Grams)	2511.57	1885.13	1433.58
	(848.43)	(1048.06)	(845.84)
Apgar > 7 at 5 Min. (%)	88.50	70.00	64.17
	(31.90)	(45.90)	(48.05)
C-Section (%)	40.68	55.67	64.17
	(49.13)	(49.76)	(48.05)
Diagnostic-Related Group (Severity)	2.19	3.24	3.43
	(0.88)	(0.88)	(0.72)
Diagnostic-Related Group (% Surgical)	6.97	34.67	33.75
	(25.47)	(47.67)	(47.38)
First Birth (%)	74.21	79.33	86.67
	(43.75)	(40.55)	(34.06)
Length of stay (Days)	18.98	64.99	75.30
	(19.32)	(50.61)	(48.91)
Overtime Reform	16.73	22.08	16.00
	(37.20)	(40.71)	(35.48)
	NICU (1,822 DAYS)		
Daily Admissions	4.37	4.33	4.55
	(2.29)	(2.30)	(2.36)
Bed Occupancy	50.37	50.72	50.71
	(3.33)	(3.48)	(3.07)
Daily Regular Hours	519.82	535.15	522.35
	(49.92)	(45.46)	(50.50)
Daily Overtime Hours	22.70	26.77	26.53
	(20.74)	(21.62)	(21.84)
	OUTCOMES		
Number of Events (Infants)	3,513	389 [300][†]	272 [240][†]
Infant Frequency (%)		7.54	6.03
Daily Frequency (%)		21.35	14.92
Daily/Infant Frequency (%)		0.458	0.321

[†] The number between brackets represents the number of neonates involved in the events.

Table 2: MFVB Estimation Results: Probability of MA

	Coeff.	Sd	Odds ratio	Sd	min 95% CI	max 95% CI
Constant coefficients: β^R						
Intercept	-7.50897	0.29623				
Birth Characteristics:						
Prior Infection	-0.51989	0.03463	0.59459	0.02059	0.55557	0.63634
Gestational Age	0.05979	0.00669	1.06161	0.00710	1.04778	1.07563
Birth vs Transfer	-0.04083	0.03693	0.95999	0.03545	0.89297	1.03205
Sex (Female=1)	0.02381	0.02342	1.02409	0.02398	0.97815	1.07219
C-Section	-0.07949	0.02967	0.92359	0.02740	0.87141	0.97889
Twins	-0.06311	0.03019	0.93884	0.02834	0.88490	0.99607
Apgar Score	-0.26775	0.02968	0.76510	0.02271	0.72186	0.81093
Surgical (vs Medical) Admission	0.34348	0.02979	1.40985	0.04200	1.32988	1.49462
DRG Severity (Omitted: Very High)						
Low	-1.04717	0.05334	0.35093	0.01872	0.31610	0.38961
Medium	-1.09140	0.03689	0.33575	0.01239	0.31233	0.36092
High	-0.38154	0.03155	0.68281	0.02155	0.64186	0.72637
NICU Characteristics:						
Occupancy	-0.03526	0.00422	0.96535	0.00408	0.95739	0.97338
Daily Admissions	-0.00305	0.00529	0.99696	0.00528	0.98667	1.00736
Overtime Reform (June 2012)	0.25988	0.03390	1.29678	0.04397	1.21340	1.38588
Constant coefficients β^G of semi-parametric function $f(X^G)$:						
Daily Overtime Hours	0.00708	0.00058				
Daily Regular Hours	0.00492	0.00026				
Birth Weight	-0.00012	0.00004				
$\sigma^2_{uR} \times 10^3$	0.00015					
$\rho_{\sigma^2_{uR}} \times 10^3$	0.00005					

Notes:

1. $\rho_{\sigma^2_{uR}}$ is the proportion of the total variance contributed by the panel-level variance component (σ^2_{uR}): $\rho_{\sigma^2_{uR}} = \sigma^2_{uR}/(\sigma^2_{uR} + \sigma^2_\epsilon)$ where $\sigma^2_\epsilon = \pi^2/3$ is the variance of the logistic distribution.

2. Hours of work are contemporaneous to y_{it}

Table 3: MFVB Results: Probability of HCAI

|	Coeff.	Sd	Odds ratio	Sd	min 95% CI	max 95% CI
Constant coefficients: β^R						
Intercept	-4.87680	0.30249				
Birth Characteristics:						
Prior Accident	-0.46634	0.03171	0.62729	0.01989	0.58949	0.66752
Gestational Age	-0.01492	0.00698	0.98519	0.00688	0.97181	0.99876
Birth vs Transfer	0.27763	0.03910	1.31999	0.05162	1.22260	1.42514
Sex (Female=1)	-0.01500	0.02427	0.98512	0.02390	0.93936	1.03310
C-section	-0.20297	0.03066	0.81631	0.02503	0.76870	0.86686
Twins	0.05863	0.03064	1.06039	0.03249	0.99858	1.12602
Apgar Score	-0.03327	0.03035	0.96727	0.02935	0.91142	1.02655
Surgical (vs Medical) Admission	0.30544	0.03098	1.35723	0.04204	1.27727	1.44218
DRG Severity (Omitted: Very High)						
Low	-2.31153	0.05936	0.09911	0.00588	0.08822	0.11134
Medium	-1.13585	0.03878	0.32115	0.01245	0.29765	0.34651
High	-0.17249	0.03265	0.84157	0.02748	0.78940	0.89718
NICU Characteristics:						
Occupancy	-0.00313	0.00438	0.99688	0.00437	0.98835	1.00548
Daily Admissions	0.03458	0.00547	1.03519	0.00566	1.02415	1.04635
Overtime Reform (June 2012)	-0.00098	0.03598	0.99902	0.03595	0.93100	1.07202
Constant coefficients β^G of semi-parametric function $f(X^G)$:						
Daily Overtime Hours[†]	0.00631	0.00060				
Daily Regular Hours[†]	-0.00002	0.00027				
Birth Weight	-0.00023	0.00004				
$\sigma^2_{u_R} \times 10^3$	0.00009					
$\rho_{\sigma^2_{u_R}} \times 10^3$	0.00003					

Notes:

1. [†]Hours of work are computed as 3-day moving averages.
2. $\rho_{\sigma^2_{u_R}}$ is the proportion of the total variance contributed by the panel-level variance component
   ($\sigma^2_{u_R}$): $\rho_{\sigma^2_{u_R}} = \sigma^2_{u_R}/(\sigma^2_{u_R} + \sigma^2_\epsilon)$ where $\sigma^2_\epsilon = \pi^2/3$ is the variance of the logistic distribution.
3. Hours of work are contemporaneous to y_{it}

Figure 3: Contour Plots of Odds-Ratios for Medical Accidents.
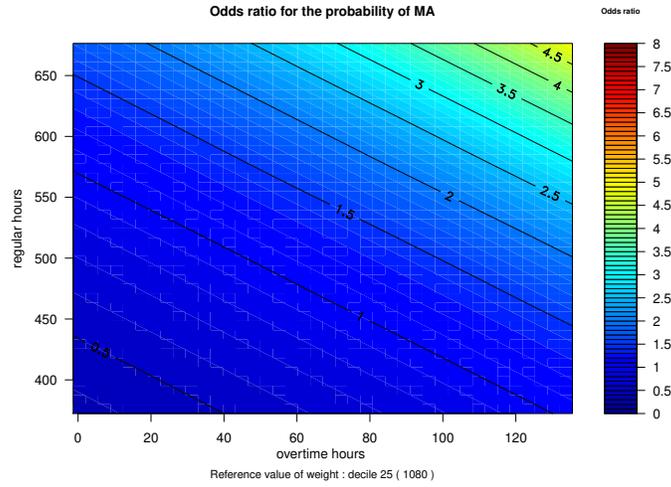
(a) Regular *vs* Overtime Hours (Birth Weight: Decile 25)
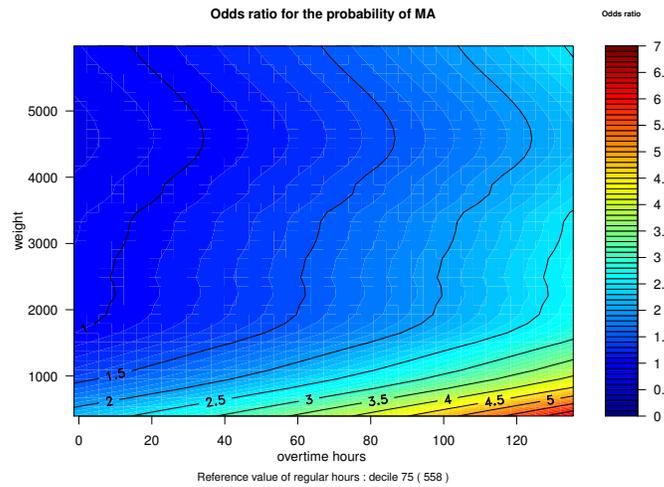
(b) Birth Weight *vs* Overtime Hours (Regular Rours: Decile 75)
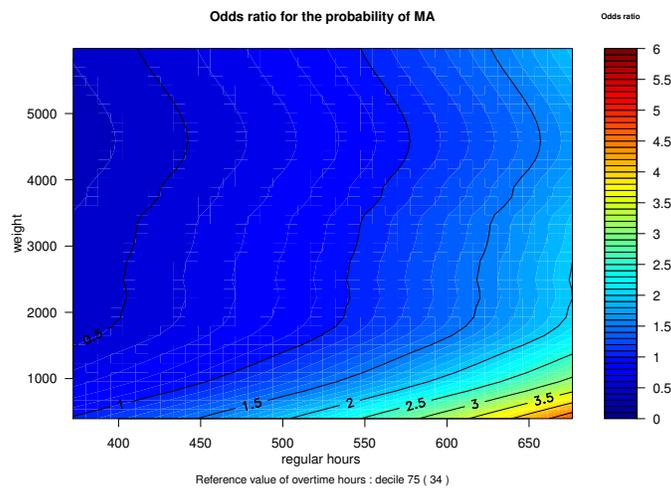

(c) Birth Weight *vs* Regular Hours (Overtime Hours: Decile 75)

Figure 4: Contour Plots of Odds-Ratios for Health Care Associated Infections.
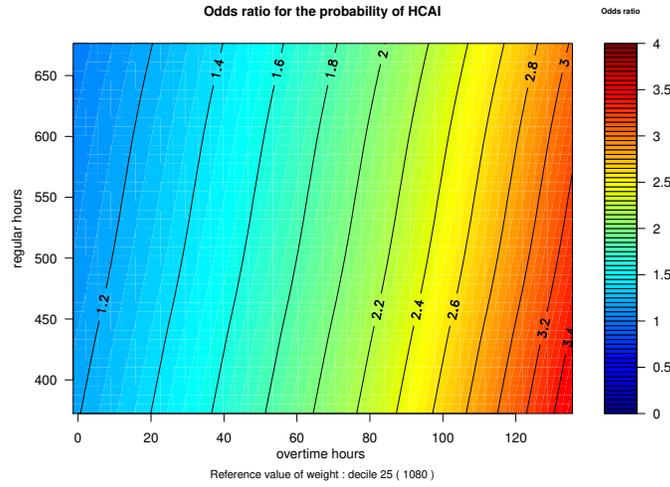
(a) Regular and Overtime hours (Birth Weight: Decile 25)
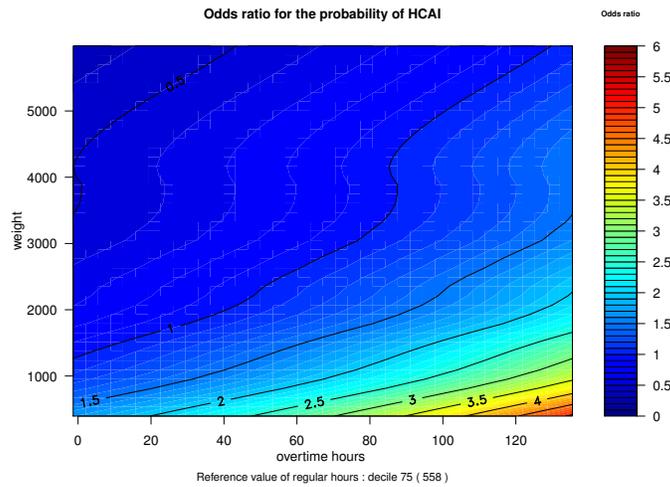
(b) Birth Weight *vs* Overtime Hours (Regular Hours: Decile 75)
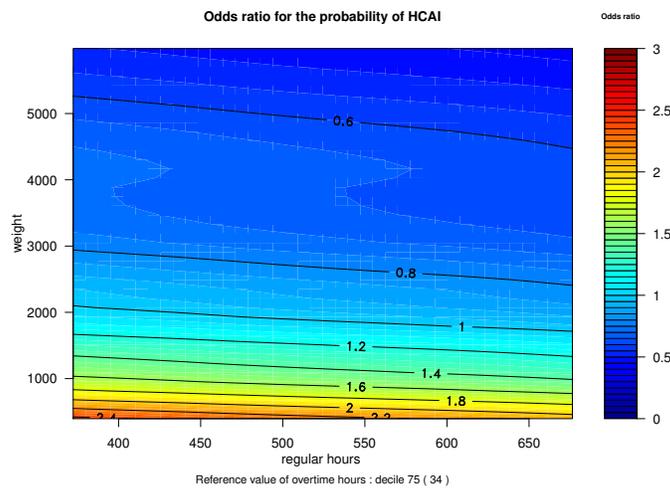
(c) Birth Weight *vs* Regular Hours (Overtime Hours: Decile 75)

References

Bae, S.H., 2013. Presence of nurse mandatory overtime regulations and nurse and patient outcomes. Nursing Economic$ 31, 59–89.

Bae, S.H., Favry, D., 2013. Assessing the relationships between nurse work hours/overtime and nurse and patient outcomes: Systematic literature review. Nursing Outlook 62, 138–156.

Behrman, R.E., Butler, A.S., 2007. Preterm Birth: Causes, Consequences, and Prevention. The National Academies Press.

Beltampo, M., Bresson, G., Lacroix, G., 2019. Using machine learning to predict nosocomial infections and medical accidents in a NICU. British Medical Journal, ADC Fetal & Neonatal (submitted) .

Beltempo, M., Blais, R., Lacroix, G., Cabot, M., Piedboeuf, B., 2017. Association of nursing overtime, nurse staffing, and unit occupancy with health care–associated infections in the NICU. American Journal of Perinatology 34, 996–1002.

Beltempo, M., Lacroix, G., Cabot, M., Piedboeuf, B., 2016. Factors and costs associated with the use of registered nurse overtime in the neonatal intensive care unit. Pediatrics and Neonatal Nursing Open Journal 4, 17–23.

Berney, B., Needleman, J., 2005. Trends in nurse overtime, 1995-2002. Policy Polit Nurs Pract 6, 183–90.

Bharadwaj, P., Lundborg, P., Rooth, D.O., 2018. Birth weight in the long run. Journal Human Resources 53, 189–231.

Bishop, C., 2006. Pattern Recognition and Machine Learning. Springer, New York, USA.

Black, S.E., Devereux, P.J., Salvanes, K.G., 2007. From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes*. The Quarterly Journal of Economics 122, 409–439.

Blei, D., Kucukelbir, A., McAuliffe, J., 2017. Variational inference: A review for statisticians. Journal of the American Statistical Association 112, 518, 859–877.

Canadian Association of Paediatric Health Care Centers, 2013. Benchmarking report 2013.

Carroll, A.E., 2015. The Concern for Supply-Sensitive Neonatal Intensive Care Unit Care: If You Build Them, They Will Come. JAMA Pediatrics 169, 812–813.

Cimiotti, J., Haas, J., Saiman, L., Larson, E., 2006. Impact of staffing on bloodstream infections in the neonatal intensive care unit. Archive of Pediatrics and Adolescent Medicine 160, 832–836.

Cimiotti, J.P., Aiken, L.H., Sloane, D.M., Wu, E.S., 2012. Nurse staffing, burnout, and health care-associated infection. American Journal of Infection Control 40, 486–490.

Cook, A., Gaynor, M., Stephens Jr, M., Taylor, L., 2012. The effect of a hospital nurse staffing mandate on patient health outcomes: Evidence from California's minimum staffing regulation. Journal of Health Economics 31, 340–348.

Currie, J., Stabile, M., Manivong, P., Roos, L.L., 2010. Child health and young adult outcomes. The Journal of Human Resources 45, 517–548.

Dorrian, J., Lamond, N., van den Heuvel, C., Pincombe, J., Rogers, A.E., Dawson, D., 2006. A pilot study of the safety implications of Australian nurses' sleep and work hours. Chronobiology International 23, 1149–1163.

Evans, W.N., Kim, B., 2006. Patient outcomes when hospitals experience a surge in admissions. Journal of Health Economics 25, 365–388.

Faes, C., Ormerod, J., Wand, M., 2011. Variational Bayesian inference for parametric and non-parametric regression with missing predictor data. Journal of the American Statistical Association 106, 959–971.

Fallah, S., Chen, X.K., Lefebvre, D., Kurji, J., Hader, J., Leeb, K., 2011. Babies admitted to nicu/icu: Province of birth and mode of delivery matter. Healthcare Quarterly 14, 16–20.

Figlio, D., Guryan, J., Karbownik, K., Roth, J., 2014. The effects of poor neonatal health on children's cognitive development. American Economic Review 104, 3921–55.

Figueiras, A., Cadarso-Suárez, C., 2012. Application of nonparametric models for calculating odds ratios and their confidence intervals for continuous exposures. American Journal of Epidemiology 154, 3, 264–275.

Freedman, S., 2016. Capacity and utilization in health care: The effect of empty beds on neonatal intensive care admission. American Economic Journal: Economic Policy 8, 154–85.

Freeman, J., Epstein, M.F., Smith, N.E., Platt, R., Sidebottom, D.G., Goldmann, D.A., 1990. Extra Hospital Stay and Antibiotic Usage With Nosocomial Coagulase-Negative Staphylococcal Bacteremia in Two Neonatal Intensive Care Unit Populations. American Journal of Diseases of Children 144, 324–329.

Goodman, D.C., Fisher, E.S., Little, G.A., Stukel, T.A., Chang, C.h., Schoendorf, K.S., 2002. The relation between the availability of neonatal intensive care and neonatal mortality. New England Journal of Medicine 346, 1538–1544.

Griffiths, P., Dall'Ora, C., Simon, M., Ball, J., Lindqvist, R., Rafferty, A.M., et al., 2014. Nurses' shift length and overtime working in 12 european countries: the association with perceived quality of care and patient safety. Med Care 52, 975–981.

Grumbach, K., 2002. Specialists, technology, and newborns — too much of a good thing. New England Journal of Medicine 346, 1574–1575.

Haizhen, L., 2014. Revisiting the relationship between nurse staffing and quality of care in nursing homes: An instrumental variables approach. Journal of Health Economics 37, 13 – 24.

Harrison, W., Goodman, D., 2015. Epidemiologic Trends in Neonatal Intensive Care, 2007-2012. JAMA Pediatrics 169, 855–862.

Honoré, B., Lewbel, A., 2002. Semiparametric binary choice panel data without strictly exogenous regressors. Econometrica 68, 2053–2063.

Hugonnet, S., Harbartha, S., Saxa, H., Duncanb, R.A., Pitteta, D., 2006. Nursing resources: A major determinant of nosocomial infection ? Archives of Pediatrics and Adolescent Medicine 160, 832–836.

Lee, C.Y.Y., 2016. Fast Approximate Inference for Longitudinal and Multilevel Data Analysis. Ph.D. thesis. University of Technology Sydney, Australia.

Lee, C.Y.Y., Wand, M., 2016. Streamlined mean field variational bayes for longitudinal and multilevel data analysis. Biometrical Journal 58, 4, 868–895.

Lewbel, A., 2000. Semiparametric qualitative response model estimation with unknown heteroskedasticity or instrumental variables. Journal of Econometrics 97, 145–177.

Mistry, H., Dowie, R., Franklin, R.C., Jani, B.R., 2009. Costs of neonatal care for low-birthweight babies in english hospitals. Acta Paediatrica 98, 1123–1129.

Moon, H., Shum, M., Weidner, M., 2017. Estimation of random coefficients Logit demand models with interactive fixed effects. Working Paper No. 17-10. USC Dornsife Institute for New Economic Thinking, CA, USA.

Mújica-Mota, R.E., Landa, P., Pitt, M., Allen, M., Spencer, A., 2020. The heterogeneous causal effects of neonatal care: a model of endogenous demand for multiple treatment options based on geographical access to care. Health Economics 29, 46–60.

Neelon, B., Anthopolos, R., Miranda, M.L., 2014. A spatial bivariate probit model for correlated binary data with application to adverse birth outcomes. Statistical Methods in Medical Research 23, 119–133.

Oreopoulos, P., Stabile, M., Walld, R., Roos, L.L., 2008. Short-, medium-, and long-term consequences of poor infant health an analysis using siblings and twins. Journal Human Resources 43, 88–138.

Ormerod, J., Wand, M., 2010. Explaining variational approximations. The American Statistician 64, 2, 140–153.

Paneth, N.S., 1995. The problem of low birth weight. The Future of Children 5, 19–34.

Pham, T., Ormerod, J., Wand, M., 2013. Mean field variational Bayesian inference for nonparametric regression with measurement error. Computational Statistics and Data Analysis 68, 375–387.

Polin, R.A., Denson, S., Brady, M.T., 2012. Strategies for prevention of health care associated infections in the nicu. Pediatrics 129, 1085–1093.

Roemer, M.I., 1961. Bed supply and hospital utilization: a natural experiment. Hospitals 35, 36.

Rogowski, J.A., Staiger, D., Patrick, T., Horbar, J., Kenny, M., Lake, E.T., 2013. Nurse Staffing and NICU Infection Rates. JAMA Pediatrics 167, 444–450.

Ruppert, D., Wand, M., Caroll, R., 2003. Semiparametric Regression. Cambridge University Press, New York, USA.

Russell, R.B., Green, N.S., Steiner, C.A., Meikle, S., Howse, J.L., Poschman, K., Dias, T., Potetz, L., Davidoff, M.J., Damus, K., Petrini, J.R., 2007. Cost of hospitalization for preterm and low birth weight infants in the United States. Pediatrics 120, 1–9.

Sherenian, M., Profit, J., Schmidt, B., Suh, S., Xiao, R., Zupancic, J.A.F., DeMauro, S.B., 2013. Nurse-to-patient ratios and neonatal outcomes: A brief systematic review. Neonatology 104, 179–183.

Sochalski, J., Konetzka, R.T., Zhu, J., Volpp, K., 2008. Will mandated minimum nurse staffing ratios lead to better patient outcomes? Medical Care 46, 606–613.

Trinkoff, A.M., Johantgen, M., Storr, C.L., Gurses, A.P., Liang, Y., Han, K., 2011. Nurses' work schedule characteristics, nurse staffing, and patient mortality. Nursing Research 60, 1–8.

Tucker, J., Tarnow-Mordi, W., Gould, C., Parry, G., Marlow, N., 1999. On behalf of the UK neonatal staffing study collaborative group. uk neonatal intensive care services in 1996. Child Fetal Neonatal Ed 80, 233–34.

Vain, N., Fariña, D., Vázquez, L., 2012. Neonatology in the emerging countries: the strategies and health-economics challenges related to prevention of neonatal and infant infections. Early Human Development 88, S53 – S59. Proceedings and Selected abstracts of the 3rd International Conference on Clinical Neonatology.

Wand, M., Ormerod, J., 2008. On semiparametric regression with O'sullivan penalized splines. Australian and New Zealand Journal of Statistics 50, 179–198.

Watson, S.I., Arulampalam, W., Petrou, S., Marlow, N., Morgan, A.S., Draper, E., Modi, N., 2016. The effects of a one-to-one nurse-to-patient ratio on the mortality rate in neonatal intensive care: a retrospective, longitudinal, population-based study. Archives of Disease in Childhood - Fetal and Neonatal Edition 101, F195–F200.

Weinstein, R.A., Stone, P.W., Pogorzelska, M., Kunches, L., Hirschhorn, L.R., 2008. Hospital staffing and health care-associated infections: A systematic review of the literature. Clinical Infectious Diseases 47, 937–944.

Zao, Y., Staudenmayer, J., Coull, B., Wand, M., 2006. General design Bayesian generalized linear mixed models. Statistical Science 21, 35–51.

Appendix A. The mean field variational Bayes approximation

Lee and Wand (2016) derived the full Bayesian model (with priors on parameters and hyperparameters). Let

$$y \mid \beta, u \sim \text{Bernoulli}\left(\text{logit}^{-1}\left(X^R \beta^R + Z^R u^R + X^G \beta^G + Z^G u^G\right)\right) \qquad (A.1)$$

with

$$
\begin{cases}
\beta & \sim & N\left(0, \sigma_\beta^2 I_P\right) \\
y \mid \Sigma^R, \sigma_{u_l}^2 & \sim & N\left(0, \begin{pmatrix} I_N \otimes \Sigma^R & 0 \\ 0 & \underset{(1 \le l \le L)}{\text{blockdiag}}\left(\sigma_{u_l}^2 I_{q_l^G}\right) \end{pmatrix}\right) \\
\Sigma^R \mid a_1^R, .., a_{q^R}^R & \sim & IW\left(\nu + q^R - 1, 2\nu \text{diag}\left(1/a_1^R, .., 1/a_{q^R}^R\right)\right) \\
a_r^R & \sim & IG\left(\frac{1}{2}, A_{R_r}^{-2}\right), 1 \le r \le R \\
\sigma_{u_l}^2 \mid a_{u_l} & \sim & IG\left(\frac{1}{2}, 1/a_{u_l}\right), 1 \le l \le L \\
a_{u_l} & \sim & IG\left(\frac{1}{2}, A_{u_l}^{-2}\right)
\end{cases},
$$

where $IW(.)$ and $IG(.)$ are inverse-Wishart and inverse-Gamma distributions.[19] The likelihood combined with the prior yields a joint posterior distribution which does not have a known tractable distribution and the parameters have to be sampled using MCMC techniques such as Gibbs sampling. Inference based on MCMC can be very slow for such models and may suffer from poor mixing. In this case, variational Bayesian inference which is a deterministic optimization approach to approximate the posterior distribution is preferred. The MFVB approximation is analogous to Gibbs sampling for conjugate models (see Bishop (2006), Ormerod and Wand (2010), Pham et al. (2013) and Lee and Wand (2016) to mention a few).

In what follows, we provide a brief overview of the MFVB method and its application to the semiparametric Logit model with random coefficients. Consider a generic Bayesian model with observed vector y and parameter vector θ. The Bayes theorem allows one to define the posterior distribution as:

$$p(\theta \mid y) = \frac{p(\theta, y)}{p(y)} = \frac{p(y \mid \theta) p(\theta)}{p(y)} \text{ with } p(y) = \int p(\theta, y) d\theta. \qquad (A.2)$$

Let $\{\theta_1, ..., \theta_M\}$ be a partition of the parameter vector θ. The MFVB approximates

---

[19] The initial values of the hyperparameters of the priors are: $\sigma_\beta^2 = 10^5$, $A_u = 10^5$, $A_R = 10^5$ and $\nu = 2$ leading to diffuse priors. We use transformed cubic O'Sullivan splines with 25 interior knots.

the posterior distribution $p(\theta \mid y)$ by the product of the q-densities[20]:

$$q(\theta) = \prod_{j=1}^{M} q(\theta_j).$$ (A.3)

Then, the logarithm of the marginal likelihood satisfies:

$$
\begin{aligned}
\ln p(y) &= \int q(\theta) \ln \left\{ \frac{p(\theta, y)}{q(\theta)} \right\} d\theta + \int q(\theta) \ln \left\{ \frac{q(\theta)}{p(\theta \mid y)} \right\} d\theta \\
&= \ln \underline{p}(y, q) + KL(q, p),
\end{aligned}
$$ (A.4)

where

$$KL(q, p) = \int q(\theta) \ln \left\{ \frac{q(\theta)}{p(\theta \mid y)} \right\} d\theta$$ (A.5)

is the Kullback-Leibler divergence between $q(\theta)$ and $p(\theta \mid y)$. Furthermore, $\ln \underline{p}(y, q)$ is a lower bound on the marginal log-likelihood.

The *optimal q*-densities which minimize the Kullback-Leibler divergence are given by:

$$q(\theta_j) \propto \exp\left[E_{q(-\theta_j)} \{ p(\theta_j \mid \text{ rest}) \} \right], j = 1, ..., M,$$ (A.6)

where $E_{q(-\theta_j)}$ denotes expectation with respect to $\prod_{k \neq j} q(\theta_k)$. rest $= \{y, \theta_1, \cdots \theta_{j-1}, \theta_{j+1} \cdots \theta_M\}$ is the set containing the rest of the random vectors in the model, except θ_j and the distributions $(\theta_j \mid \text{rest})$ are the full conditionals in the MCMC literature. The Kullback-Leibler divergence becomes

$$
\begin{aligned}
KL(q, p) &= E_{q(\theta)} \left[\ln q(\theta) \right] - E_{q(\theta)} \left[\ln p(\theta \mid y) \right] \\
&= E_{q(\theta)} \left[\ln q(\theta) \right] - E_{q(\theta)} \left[\ln p(\theta, y) \right] + \ln p(y),
\end{aligned}
$$ (A.7)

where the last term, $\ln p(y)$, is a constant. The minimization of the Kullback-Leibler divergence is thus equivalent to maximizing the scalar quantity,

$$\ln \underline{p}(y, q) = E_{q(\theta)} \left[\ln \left(\frac{p(\theta, y)}{q(\theta)} \right) \right]$$ (A.8)

which is usually referred as the evidence lower bound (ELBO)[21]. Compared to the

---

[20]This is known as the *mean field restriction*. The term *mean field* originated from physics.

[21]The lower bound is also known as the negative variational free energy and the entropy of the variational distribution is given by $E_{q(\theta)} \log [q(\theta)]$.

minimization of the KL divergence, the maximization of the ELBO is often a more convenient objective of the optimization over the free distributional parameters.

Lee and Wand (2016) apply this principle and derive the MFVB approximation of the logistic-mixed model-based penalized spline specification (A.1) on the following factorization:

$$
\begin{aligned}
p\left(\theta \mid y\right) &= p\left(\theta_1, ..., \theta_M \mid y\right) = p\left(\beta, u, \Sigma^R, a_u, a^R, \sigma_u^2 \mid y\right) & \text{(A.9)} \\
&= p\left(\beta^R, \beta^G, u^R, u^G, \Sigma^R, a_u, a^R, \sigma_u^2 \mid y\right) \\
&\approx q\left(\beta, u, \Sigma^R, a_u, a^R, \sigma_u^2\right) = q\left(\beta, u, a_u, a^R\right) q\left(\Sigma^R, \sigma_u^2\right) \\
&= q\left(\beta, u\right) q\left(\Sigma^R\right) \prod_{r=1}^{q^R} q\left(a_r^R\right) \prod_{l=1}^{L} q\left(a_{u_l}\right) \prod_{l=1}^{L} q\left(\sigma_{u_l}^2\right).
\end{aligned}
$$

They approximate the optimal q-density for (β, u) by a multivariate normal distribution

$$
\begin{aligned}
q\left(\beta, u\right) \quad &\sim \quad N\left(\mu_{q(\beta,u,\xi)}, \Sigma_{q(\beta,u,\xi)}\right) & \text{(A.10)} \\
&\text{with} \quad \Sigma_{q(\beta,u,\xi)} = \left[2C'diag\left(\lambda\left(\xi\right)\right) C + \text{blockdiag}\left(\sigma_\beta^{-1} I_P, G^{-1}\right)\right]^{-1} \\
&\quad \mu_{q(\beta,u,\xi)} = \Sigma_{q(\beta,u,\xi)} C'\left(y - \frac{1}{2}\mathbf{1}\right),
\end{aligned}
$$

where $C = [X, Z]$, $\lambda\left(\xi\right) = \tanh\left(\xi/2\right)/\left(4\xi\right)$, $\mathbf{1}$ is a $(NT \times 1)$ vector of ones and ξ is an $(T \times 1)$ vector of positive variational parameters.[22]

Lee and Wand (2016) first derive the conditional posterior densities $p\left(\theta_j \mid \text{rest}\right)$ for $j = 1, ..., M$ from the full Bayesian model (A.1), *i.e.*, the Gibbs sampling algorithm. Then, the optimal q-densities are chosen to minimize Kullback-Leibler divergence between the right-hand side of (A.9) and the full joint posterior density function from the Gibbs sampling algorithm. These optimal q-densities admit the

---

[22]The optimal variational parameter vector is obtained *via* the update

$$
\xi \leftarrow \sqrt{\text{diagonal}\left[C\left(\Sigma_{q(\beta,u,\xi)} + \mu_{q(\beta,u,\xi)}\mu'_{q(\beta,u,\xi)}\right)C'\right]}
$$

where diagonal(x) is the vector containing the diagonal entries of x.

following forms

$$\begin{cases} q\left(\beta, u\right) & \sim & N\left(\mu_{q(\beta,u,\xi)}, \Sigma_{q(\beta,u,\xi)}\right) \\ q\left(\Sigma^R\right) & \sim & IW\left(\nu + N + q^R - 1, B_{q(\Sigma^R)}\right) \\ q\left(\sigma^2_{u_l}\right) & \sim & IG\left(\frac{1}{2}\left(q_l^G + 1\right), B_{q(\sigma^2_{u_l})}\right) \\ q\left(a_{u_l}\right) & \sim & IG\left(1, B_{q(a_{u_l})}\right) \\ q\left(a_r^R\right) & \sim & IG\left(\frac{1}{2}\left(\nu + q^R\right), B_{q(a_r^R)}\right) \end{cases} \qquad (A.11)$$

After tedious derivations, the optimal values of the q-density parameters are obtained *via* coordinate ascent in the following Algorithm 1. The stopping criterion is based on the variational lower bound on the marginal likelihood denoted by $\underline{p}\left(y, q\right)$ and its logarithm is obtained by

$$\begin{aligned} \ln \underline{p}\left(y, q\right) & = & E_{q(\theta)}\left[\ln\left(\frac{p\left(\theta, y\right)}{q\left(\theta\right)}\right)\right] \qquad\qquad (A.12) \\ & = & E_{q(\theta)}\left[\ln p\left(y, \beta, u, \Sigma^R, a_u, a^R, \sigma^2_u\right) - \ln q\left(\beta, u, \Sigma^R, a_u, a^R, \sigma^2_u\right)\right] \end{aligned}$$

and is presented below.

The computing time gains afforded by the MFVB algorithm, as compared to Gibbs sampling, are huge[23] but more importantly, this approximation avoids the pitfalls of poor mixing of MCMC methods on models with large sparse covariance matrices. Moreover, the accuracy scores of the MFVB approximation (as compared to MCMC) generally exceed $95 - 97\%$ and rarely drop below 90% in most of the papers on MFVB (see for instance Bishop (2006), Ormerod and Wand (2010), Faes et al. (2011), Pham et al. (2013), Lee and Wand (2016) and Blei et al. (2017)).

Thus, the MFVB approach has great advantages as compared to the MCMC technique such as Gibbs sampling. The algorithm proposed by Lee and Wand (2016) has a promising future with a remarkable ability to perform high quality Bayesian inference for large panel data models faster than ever before.

---

[23]The MFVB can be $3,500$ (or more) times faster than the MCMC. For instance, Lee and Wand (2016) used a data set with $N = 3,978$ mothers as groups at the first level and $8,604$ births as units at level 2 for an application on the link between birthweight of children and infant's gestational age (in weeks). It took 4 days with MCMC and few minutes with MFVB on a Mac OS X laptop with a 2.G GHz Intel Core i5 and 8 GB RAM. They also used a data set with $N = 148$ schools as groups at the first level and $2,069$ students as units at level 2 for an application to student assessment which took 57 min with MCMC and 1.2 min with MFVB.

Algorithm 1. Mean field variational Bayes algorithm (see Lee and Wand (2016), pp. 884-885).

---

1. Initialize $\mu_{q(1/\sigma_{u_l}^2)} > 0$, $\mu_{q(1/a_{u_l})} > 0$, $1 \leq l \leq L$, $\mu_{q(1/a_r^R)} > 0$, $1 \leq r \leq q^R$, $M_{q((\Sigma^R)^{-1})}$ positive definite, ξ $(T \times 1)$ vector of positive entries.

2. Cycle through updates:

   (a) $S \leftarrow 0$, $s \leftarrow 0$, for $i = 1, ..., N$:
   $$G_i \leftarrow 2\left(C_i^G\right)' diag\left(\lambda\left(\xi_i\right)\right) X_i^R \text{ with } C_i^G = \left[X_i^G, \; Z_i^G\right]$$
   $$H_i \leftarrow \left[2\left(X_i^R\right)' diag\left(\lambda\left(\xi_i\right)\right) X_i^R + M_{q((\Sigma^R)^{-1})}\right]^{-1}$$
   $$S \leftarrow S + G_i H_i\left(G_i\right)'; \; s \leftarrow s + G_i H_i\left(X_i^R\right)'\left(y_i - \tfrac{1}{2}\mathbf{1}\right)$$

   (b) $$\Sigma_{q(\beta,u^G,\xi)} \leftarrow \left[2\left(C^G\right)' diag\left(\lambda\left(\xi\right)\right) C^G + \begin{bmatrix} \sigma_\beta^{-2} I_P & 0 \\ 0 & \text{blockdiag}\limits_{(1\leq l\leq L)}\left(\mu_{q(1/\sigma_{u_l}^2)} I_{q_l^G}\right) \end{bmatrix} - S\right]^{-1}$$

   (c) $\mu_{q(\beta,u^G,\xi)} \leftarrow \Sigma_{q(\beta,u^G,\xi)}\left[\left(C^G\right)'\left(y_i - \tfrac{1}{2}\mathbf{1}\right) - s\right]$ with $C^G = \left[X^G, \; Z^G\right]$

   (d) for $i = 1, ..., N$:
   $$\Sigma_{q\left(u_i^R,\xi\right)} \leftarrow H_i + H_i G_i' \Sigma_{q(\beta,u^G,\xi)} G_i H_i$$
   $$\mu_{q\left(u_i^R,\xi\right)} \leftarrow H_i\left[\left(X_i^R\right)'\left(y_i - \tfrac{1}{2}\mathbf{1}\right) - G_i'\mu_{q(\beta,u^G,\xi)}\right]$$

   (e) $\xi^2 \leftarrow \text{diagonal}\left[C^G\left(\Sigma_{q(\beta,u^G,\xi)} + \mu_{q(\beta,u^G,\xi)}\mu'_{q(\beta,u^G,\xi)}\right)\left(C^G\right)'\right]$

   (f) for $i = 1, ..., N$:
   $$\xi^2 \leftarrow \xi^2 + 2\,\text{diagonal}\left[C^G\left(-\Sigma_{q(\beta,u^G,\xi)} G_i H_i + \mu_{q(\beta,u^G,\xi)}\mu'_{q(\beta,u^G,\xi)}\right)\left(X_i^R\right)'\right]$$
   $$\xi^2 \leftarrow \xi^2 + 2\,\text{diagonal}\left[X_i^R\left(\Sigma_{q\left(u_i^R,\xi\right)} + \mu_{q\left(u_i^R,\xi\right)}\mu'_{q\left(u_i^R,\xi\right)}\right)\left(X_i^R\right)'\right]$$

   (g) for $r = 1, ..., q^R$:
   $$B_{q(a_r^R)} \leftarrow \nu\left(M_{q((\Sigma^R)^{-1})}\right)_{rr} + A_{Rr}^{-2}; \; \mu_{q(1/a_r^R)} \leftarrow \tfrac{1}{2}\left(\nu + q^R\right)/B_{q(a_r^R)}$$

   (h) $B_{q(\Sigma^R)} \leftarrow \sum_{i=1}^N\left(\mu_{q\left(u_i^R,\xi\right)}\mu'_{q\left(u_i^R,\xi\right)} + \Sigma_{q\left(u_i^R\right)}\right) + 2\nu\,\text{diag}\left(\mu_{q(1/a_1^R)}, ..., \mu_{q\left(1/a_{q^R}^R\right)}\right)$

   (i) $M_{q((\Sigma^R)^{-1})} \leftarrow \left(\nu + N + q^R - 1\right) B_{q(\Sigma^R)}^{-1}$

   (j) for $l = 1, ..., L$:
   $$\mu_{q(1/a_{u_l})} \leftarrow 1/\left[\mu_{q(1/\sigma_{u_l}^2)} + A_{u_l}^{-2}\right]$$
   $$\mu_{q(1/\sigma_{u_l}^2)} \leftarrow \frac{q^G + 1}{2\mu_{q(1/a_{u_l})} + \|\mu_{q\left(u_l^G,\xi\right)}\|^2 + \text{tr}\left[\Sigma_{q\left(u_l^G,\xi\right)}\right]}$$

until the increase in the ELBO $\ln \underline{p}(y, q)$ is negligible.

3. for $i = 1, ..., M$:

$$\Lambda_{q\left(\beta, u^G, u_i^R, \xi\right)} \equiv E_q\left[\left(\begin{bmatrix} \beta \\ u^G \end{bmatrix} - \mu_{q(\beta, u^G, \xi)}\right)\left(u_i^R - \mu_{q\left(u_i^R, \xi\right)}\right)'\right] \leftarrow -\Sigma_{q(\beta, u^G, \xi)} G_i H_i$$

Convergence of such an algorithm to at least a local optimum is guaranteed based on convexity properties. The ELBO is judged to cease increasing when the tolerance criterion is less than 10^{-7}. This algorithm is part of the family of coordinate ascent variational inference (CAVI). It iteratively optimizes each factor of the mean field variational density, while holding the others fixed (see Bishop (2006) and Blei et al. (2017)).

The variational lower bound on the marginal log-likelihood has the following expression (see Lee and Wand (2016), pp. 893-894):

$$
\begin{aligned}
\ln \underline{p}(y, q) \;=\;& \frac{1}{2}q^R\left(\nu + q^R - 1\right)\ln 2\nu - \left(\frac{1}{2}q^R + L\right)\ln 2\pi - \lambda'(\xi)\xi^2 + \mathbf{1}'\zeta(\xi) \\
&+ \left(y - \frac{1}{2}\mathbf{1}\right)\left\{C^G\mu_{q(\beta, u^G, \xi)} + \begin{bmatrix} X_1^R\mu_{q\left(u_1^R, \xi\right)} \\ \vdots \\ X_N^R\mu_{q\left(u_N^R, \xi\right)} \end{bmatrix}\right\} + \\
&- \frac{P}{2}\ln\sigma_\beta^2 - \frac{\sigma_\beta^{-2}}{2}\left[\|\mu_{q(\beta, \xi)}\|^2 + \mathrm{tr}\left[\Sigma_{q(\beta, \xi)}\right]\right] + \frac{1}{2}\left(\sum_{l=1}^L q_l^G + P + N\right) \\
&- \frac{1}{2}\sum_{i=1}^N \ln\left|\left(X_i^R\right)' X_i^R + M_{q\left(\left(\Sigma^R\right)^{-1}\right)}\right| - \frac{1}{2}\ln\left|\Sigma_{q(\beta, u^G, \xi)}^{-1}\right| \\
&- \ln\left(Q_{q^R, \nu + q^R - 1}\right) + \ln\left(Q_{q^R, \nu + N + q^R - 1}\right) - \frac{1}{2}\left(\nu + N + q^R - 1\right)\ln\left|B_{q(\Sigma^R)}\right| \\
&+ \sum_{l=1}^L \ln\Gamma\left(\frac{q_l^G + 1}{2}\right) - \frac{1}{2}\sum_{l=1}^L\left(q_l^G + 1\right)\ln B_{q\left(\sigma_{u_l}^2\right)} - \sum_{r=1}^{q^R}\ln A_{R_r} \\
&+ q^R\ln\Gamma\left(\frac{q^R + \nu}{2}\right) + \sum_{r=1}^{q^R}\nu M_{q\left(\left(\Sigma^R\right)^{-1}\right)}\mu_{q\left(1/a_r^R\right)} - \frac{1}{2}\left(q^R + \nu\right)\sum_{r=1}^{q^R}\ln B_{q(a_r^R)} \\
&- \sum_{l=1}^L\left[\ln A_{u_l} + \ln B_{q\left(a_{u_l}\right)} - \mu_{q\left(1/a_{u_l}\right)}\mu_{q\left(1/\sigma_{u_l}^2\right)}\right]
\end{aligned}
$$

where $Q_{a,b} = 2^{ab/2}\pi^{a(a-1)/4}\prod_{j=1}^a \Gamma\left(\frac{b+1-j}{2}\right)$, $\Gamma(.)$ is the gamma function and $\zeta(x) = x/2 - \ln(1 + e^x) + x\tanh(x/2)/4$.

40

Appendix B. Odds ratios and the 95% confidence intervals

Estimated means and variances of coefficients β^R and β^G in Tables 2 and 3 come from the optimal q-density $q(\beta, u) \sim N\left(\mu_{q(\beta,u,\xi)}, \Sigma_{q(\beta,u,\xi)}\right)$ in (A.10) estimated by 2.(c) and 2.(b) in Algorithm 1. For the X_j^R covariates, its estimated coefficient $\hat{\beta}_j$ is the j-th element of $\mu_{q(\beta,u,\xi)}$ and its variance $\hat{\sigma}_{\beta_j}^2$ is the j-th element of the diagonal of $\Sigma_{q(\beta,u,\xi)}$. Let us consider the odds ratio when the X_j^R covariate changes by an amount δ, while the rest of the explanatory variables remain the same. Then, the odds ratio and its standard deviation are given by $OR\left(X_j^R\right)_\delta = \exp\left(\delta\hat{\beta}_j\right)$, $\sigma_{OR\left(X_j^R\right)_\delta} = \hat{\sigma}_{\beta_j} \exp\left(\delta\hat{\beta}_j\right)$. For dummy variables, generally $\delta = b - a$ with $a = 0$ and $b = 1$ while a and b could be specific values as deciles for a continuous variable.

For the X^G covariates (`overtime hours`, `regular hours` and `birth weight`) entering the semiparametric additive function

$$f\left(X^G\right) = X^G\beta^G + Z^G u^G = \sum_{l=1}^{L=3} X_l^G\beta_l^G + \sum_{l=1}^{L=3} Z_l^G u_l^G$$

the odds ratio of the X_l^G covariate is given by

$$OR\left(X_l^G\right) = \exp\left(f\left(X_l^G\right) - f\left(X_{l,ref}^G\right)\right) \tag{B.1}$$

with

$$f\left(X_l^G\right) = X_l^G\beta_l^G + Z_l^G u_l^G + \sum_{m(\neq l)=1}^{L=3}\left[X_{m,ref}^G\beta_m^G + Z_{m,ref}^G u_m^G\right], \tag{B.2}$$

$$f\left(X_{l,ref}^G\right) = X_{l,ref}^G\beta_l^G + Z_{l,ref}^G u_l^G + \sum_{m(\neq l)=1}^{L=3}\left[X_{m,ref}^G\beta_m^G + Z_{m,ref}^G u_m^G\right],$$

where $X_{l,ref}$ is a specific value of the exposure X_l, taken as the reference and $Z_{l,ref}$ is the associated spline basis function (see also Figueiras and Cadarso-Suárez (2012)). In our case, we choose several percentiles 10, 25, 50 and 90.

As β^G and u^G come from the optimal q-density $q(\beta, u) \sim N\left(\mu_{q(\beta,u,\xi)}, \Sigma_{q(\beta,u,\xi)}\right)$ in (A.10), we can use this distribution to draw estimates $\hat{\beta}_{Boot}^G$ and $\hat{u}_{Boot}^G$ for bootstraps

and compute[24]

$$OR\left(X_l^G\right)_{Boot} = \exp\left(f\left(X_l^G\right)_{Boot} - f\left(X_{l,ref}^G\right)_{Boot}\right), \; Boot = 1,....N_{Boot}(= 1000) \quad (B.3)$$

$$\text{with } f\left(X_l^G\right)_{Boot} = X_l^G \hat{\beta}_{l,Boot}^G + Z_l^G \hat{u}_{l,Boot}^G + \sum_{m(\neq l)=1}^{L=3}\left[X_{m,ref}^G \hat{\beta}_{m,Boot}^G + Z_{m,ref}^G \hat{u}_{m,Boot}^G\right],$$

$$\text{and } f\left(X_{l,ref}^G\right)_{Boot} = X_{l,ref}^G \hat{\beta}_{l,Boot}^G + Z_{l,ref}^G \hat{u}_{l,Boot}^G + \sum_{m(\neq l)=1}^{L=3}\left[X_{m,ref}^G \hat{\beta}_{m,Boot}^G + Z_{m,ref}^G \hat{u}_{m,Boot}^G\right].$$

The average odds ratio and its standard error allow to define a $(1-\alpha)\%$ confidence interval

$$\overline{OR}\left(X_l^G\right) \pm z_{\alpha/2} SE\left(OR\left(X_l^G\right)\right) \quad (B.4)$$

with

$$\overline{OR}\left(X_l^G\right) = \frac{1}{N_{Boot}} \sum_{Boot=1}^{N_{Boot}} OR\left(X_l^G\right)_{Boot} \quad (B.5)$$

$$SE\left(OR\left(X_l^G\right)\right) = \sqrt{\frac{1}{N_{Boot}-1} \sum_{Boot=1}^{N_{Boot}} \left(OR\left(X_l^G\right)_{Boot} - \overline{OR}\left(X_l^G\right)\right)^2}$$

For the 3D odds ratios surface, we compute

$$OR\left(X_l^G, X_m^G\right) = \exp\left(\left\{f\left(X_l^G\right) - f\left(X_{l,ref}^G\right)\right\} + \left\{f\left(X_m^G\right) - f\left(X_{m,ref}^G\right)\right\}\right) \quad (B.6)$$

$$= \exp\left[\begin{array}{c}\left(X_l^G - X_{l,ref}^G\right)\beta_l^G + \left(X_m^G - X_{m,ref}^G\right)\beta_m^G \\ +\left(Z_l^G - Z_{l,ref}^G\right)u_l^G + \left(Z_m^G - Z_{m,ref}^G\right)u_m^G\end{array}\right] \text{ with } m(\neq l) = 1 \text{ or } 2 \text{ or } 3.$$

---

[24]We do not use resampling with (or without) replacement bootstrap techniques because it is prohibitively time consuming.

Infections, Accidents and Nursing Overtime in a Neonatal Intensive Care Unit: A Bayesian Semiparametric Panel Data Logit Model

Online supplement

Marc Beltempo[a], Georges Bresson[b], Jean-Michel Etienne[c], Guy Lacroix[d,]

[a]*Department of Pediatrics, McGill University Health Centre, Montreal, QC, Canada*
[b]*Department of Economics, Université Paris II, Paris, France*
[c]*Department of Economics, Université Paris-Sud*
[d]*Department of Economics, Université Laval, Québec, QC, Canada*

Email addresses: `marc.beltempo@mcgill.ca` (Marc Beltempo), `georges.bresson@u-paris2.fr` (Georges Bresson), `jean-michel.etienne@u-psud.fr` (Jean-Michel Etienne), `Guy.Lacroix@ecn.ulaval.ca` (Guy Lacroix)

Contents

1. Wald tests on a MFVB semiparametric random intercept and slope model: a Monte Carlo simulation study

In this section, we conduct a small Monte Carlo simulation study to assess the performance of a Wald test on the random part of the random slope coefficients using the MFVB algorithm. The following simulation setting corresponds to the case of the estimation of a Bayesian semiparametric random intercept and slope model (RCM):

$$y \mid \beta, u \sim \text{Bernoulli}\left(\text{logit}^{-1}\left(X^R \beta^R + Z^R u^R + f\left(X^G\right) \right) \right) \tag{1}$$

where the notation $y \sim \text{Bernoulli}(p)$ is shorthand for the entries of y having independent Bernoulli distributions with parameters corresponding to the entries of p and $\text{logit}^{-1}(x)$ is shorthand for the logistic distribution $e^x/(1+e^x)$. X^R is an $(NT \times q^R)$ matrix of covariates — with $q^R = 2$ in this simulation study —, Z^R is an $(NT \times Nq^R)$ block-diagonal matrix of the X_i^R submatrices. X and Z are the fixed and random effects design matrices associated with β and u, the fixed effects and random effects vectors. In this small Monte Carlo simulation study, $X_{it,1}^R = 1$, $\forall i, t$ is the intercept and $X_{it,2}^R$ is the other control covariate. The random intercept is defined by the sum $(\beta_1^R + u_{i,1}^R)$, the random slope for variable $X_{i,2}$ is the sum $(\beta_2^R + u_{i,2}^R)$. Then, the random intercept and slope model (RCM) and the random intercept and constant slope model (OWEC) are defined as

$$\text{RCM:} \quad y_{it} \mid \beta_1^R, \beta_2^R, u_{i1}^R, u_{i2}^R \quad \sim \text{Bernoulli}\left(\text{logit}^{-1}\left(\left(\beta_1^R + u_{i,1}^R\right) + \left(\beta_2^R + u_{i,2}^R\right) X_{it,2} + f\left(X_{it}^G\right) \right) \right) \tag{2}$$

$$\text{OWEC:} \quad y_{it} \mid \beta_1^R, \beta_2^R, u_{i1}^R \quad \sim \text{Bernoulli}\left(\text{logit}^{-1}\left(\left(\beta_1^R + u_{i,1}^R\right) + \beta_2^R X_{it,2} + f\left(X_{it}^G\right) \right) \right) \tag{3}$$

where[1]

$$f\left(X_{it}^G\right) \quad = \quad 1 - \left(2.3 X_{it}^G - 0.07\left(X_{it}^G\right)^2\right) - 2.6\Phi\left(X_{it}^G; 0.15, 1\right) + 0.5\phi\left(X_{it}^G; 0.8, 0.07\right),$$

$$\text{with} \quad X_{it,2}^R \sim U(0,1), \, X_{it}^G \sim U(0,1), \, , \, i = 1, ..., N, \, t = 1, ..., T_i, \, NT = \sum_i^N, T_i.$$

The true parameter values are:

$$\beta_1^R = 0.58, \, \beta_2^R = 1.89.$$

A OWEC specification is generated for $N = [25, 50, 100]$ individuals with the within group sample sizes T_i $(i = 1, ..., N)$ ranged between 10 and 20 with the previous design and equation (3). In that case, $u_{i1}^R \sim N\left(0, \sigma_{u_{i1}^R}^2\right)$ where $\sigma_{u_{i1}^R}^2 = 2.58$. The MFVB fits of RCM (equation (2)) are obtained using Algorithm 1 with the cycles stopped when the relative increase in the variational lower bound on the marginal log-likelihood (ELBO) $\log \underline{p}(y, q)$ fall below 10^{-5}. The initial values of the hyperparameters of the priors

---

[1] $\phi(x; \mu, \sigma)$ (resp. $\Phi(x; \mu, \sigma)$) is the pdf (resp. the cdf) of the normal distribution with mean μ and standard deviation σ.

are: $\sigma_\beta^2 = 10^5$, $A_u = 10^5$, $A_R = 10^5$ and $\nu = 2$ leading to diffuse priors. We use transformed cubic O'Sullivan splines with 25 interior knots.

As shown in Appendix 1, the optimal q-densities for β and u admit the following form

$$q(\beta, u) \quad \sim \quad N\left(\mu_{q(\beta, u, \xi)}, \Sigma_{q(\beta, u, \xi)}\right) \tag{4}$$

So, we can define a Wald test to test the null hypothesis $H_0 : u_{i,j}^R = 0$, $j = 2, ..., q^R$. If we suppose that the true model is the random intercept and constant slope model (OWEC) and if we estimate a random intercept and slope model (RCM), we can use a Wald test to check the null hypothesis $H_0 : u_{i,2}^R = 0$ as $q^R = 2$ in this small Monte Carlo study. The statistic is defined as[2]

$$W = \left(\widehat{u}_2^R\right)' \Sigma_{\widehat{u}_2^R}^{-1}\left(\widehat{u}_2^R\right) \tag{5}$$

where $\widehat{u}_2^R = \left(\widehat{u}_{1,2}^R, ..., \widehat{u}_{N,2}^R\right)'$ coming from $\mu_{q(\beta, u, \xi)}$ and $\Sigma_{\widehat{u}_2^R}$ is a $(N \times N)$ diagonal matrix with element $\widehat{\sigma}_{u_{i2}^R}^2$ coming from $\Sigma_{q(\beta, u, \xi)}$. Under the null, this statistic follows a χ_N^2. If $W \leq \chi_N^2$ (resp. $W > \chi_N^2$), we do not reject the null and the OWEC estimation is prefered (resp. we reject the null and the RCM estimation is prefered).

With this OWEC specification and the MFVB estimation of the RCM, we compute the Wald test over $S = 200$ trials and then compute the means over the $S = 200$ trials.

With the same DGP of a random intercept and constant slope model (OWEC), we could also estimate a random intercept and constant slope model (OWEC) to compare the ELBO coming from the estimated RCM and the one coming from the OWEC. Therefore, it is possible *a priori* to use the variational lower bound as a model selection criterion just like the BIC criterion (see Beal and Ghahramani (2003), Nott et al. (2012) or Li and Sillanpää (2013) to mention a few).[3] Estimating the marginal likelihood, which is important for many applications, is argued to embody an Occam's razor. But, as shown by Dieng et al. (2017), model selection based solely on the ELBO is inappropriate because of the possible variation in the tightness of this bound. As shown by Yao et al. (2018), purely relying on the ELBO does not

---

[2]In the general case where $q^R > 2$, the Wald test to check the null hypothesis $H_0 : u_{i,j}^R = 0$, $j = 2, ..., q^R$ is

$$W = \left(\widehat{u}^R\right)' \Sigma_{\widehat{u}^R}^{-1}\left(\widehat{u}^R\right)$$

where $\widehat{u}^R = \left(\widehat{u}_{1,2}^R, ..., \widehat{u}_{1,q^R}^R, \widehat{u}_{2,2}^R, ..., \widehat{u}_{2,q^R}^R, \cdots, \widehat{u}_{N,2}^R, ..., \widehat{u}_{N,q^R}^R\right)'$ is a $(N(q^R - 1) \times 1)$ vector and $\Sigma_{\widehat{u}^R}$ is a $(N(q^R - 1) \times N(q^R - 1))$ diagonal matrix with $((q^R - 1) \times (q^R - 1))$ sub-matrices elements $\Sigma_{\widehat{u}_i^R}$, $i = 1, ..., N$:

$$\Sigma_{\widehat{u}^R} = \begin{pmatrix} \Sigma_{\widehat{u}_1^R} & 0 & \cdots & 0 \\ 0 & \Sigma_{\widehat{u}_2^R} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \Sigma_{\widehat{u}_N^R} \end{pmatrix}$$

Under the null, this statistic follows a $\chi_{N(q^R-1)}^2$.

[3]As the model with the lowest BIC is preferred, the model with the highest ELBO could be preferred *a priori*.

solve the problem. An unknown multiplicative constant exists in $p(\theta, y) \propto p(\theta \mid y)$ that changes with reparametrization, making it meaningless to compare ELBO across two approximations. Moreover, the ELBO is a quantity on an uninterpretable scale, that is it's not clear at what value of the ELBO we can begin to trust the variational posterior (see also Chérief-Abdellatif and Alquier (2018)).

For different sample sizes of (N, T_i), Table 1 gives means over 200 trials of the Wald statistics, of the p-value (or the power of the test) and the 5% empirical size[4] of the Wald test when, first, the true model is the OWEC model. Whatever the sample size $N = 25, 50$ or 100, the 200 Wald statistics of all the 200 trials are always lower than the critical values of the χ^2 distribution with p-values larger than 5%. On Table 1, the means of the 200 trials of the Wald statistics (resp. of their p-values) are small (resp. large). The 5% empirical size, corresponding to the percentage of rejections at the 5 per cent in 100 trials using Wald tests, is always lower than 0.05 whatever the number of individuals. However, we can note that this 5% empirical size is small but converge to the 5% nominal size as N increases. When the true model is a random intercept and constant slope model (OWEC), the Wald test always accepts the null hypothesis $H_0 : u_{i,2}^R = 0$ when estimating a random intercept and slope model (RCM).

In addition, Table 1 shows that the performance of the test also differs when the null hypothesis is false (*i.e.*, when the true model is a random intercept and slope model (RCM)). For that, a RCM specification has been generated for $N = [25, 50, 100]$ individuals with the within group sample sizes T_i $(i = 1, ..., N)$ ranged between 10 and 20 and with the previous design and equation (2) in which $\left[u_{i1}^R, u_{i2}^R\right]' \sim N\left(0, \Sigma^R\right)$ with

$$\Sigma^P = \left[\begin{array}{cc} \sigma_{u_{i1}^R}^2 & \sigma_{u_{i2}^R}^2/2 \\ \sigma_{u_{i2}^R}^2/2 & \sigma_{u_{i2}^R}^2 \end{array}\right] \text{ with } \sigma_{u_{i1}^R}^2 = 2.58.$$

We use several values for $\sigma_{u_{i2}^R}^2 = [0.1, 0.2, 0.5]$. With this RCM specification and the MFVB estimation of the RCM model, we compute the Wald test over $S = 200$ trials and then compute the means over the $S = 200$ trials.

If the 5% empirical size column still expresses the probability of rejecting the null (*i.e.*, reject the OWEC specification), it is not strictly speaking a real "5% empirical size" but rather a power of the test (see Gregory and Veall (1985)). Of course, the percentages are almost larger than 5%.

To conclude, when estimating an RCM model using the MFVB approximation, the Wald test clearly accepts the null hypothesis $H_0 : u_{i,2}^R = 0$ when the true model is a OWEC model and clearly rejects the null when the true model is a RCM.

---

[4]The size of the test is the probability of falsely rejecting the null hypothesis. Empirical size refers to the possibility that the nominal size that the user of the test chooses (say, 5%) may not coincide with the actual rejection frequency of the test.

Table 1: Wald tests over 200 trials.

True model	N	$\sigma^2_{u^R_{i2}}$	Wald test	$\chi^2_{N,5\%}$	p-value	empirical size
OWEC						
	25		8.5598	37.6525	0.9304	0.015
	50		20.5472	67.5048	0.9120	0.015
	100		45.6993	124.3421	0.9288	0.035
True model	N	$\sigma^2_{u^R_{i2}}$	Wald test	$\chi^2_{N,5\%}$	p-value	empirical size
RCM						
	25	0.1	13.4286	37.6525	0.8325	0.07
	25	0.2	16.5235	37.6525	0.7575	0.08
	25	0.5	25.9785	37.6525	0.5660	0.20
	50	0.1	28.1165	67.5048	0.8586	0.06
	50	0.2	35.9218	67.5048	0.7715	0.10
	50	0.5	60.5129	67.5048	0.4447	0.335
	100	0.1	58.0777	124.3421	0.8325	0.06
	100	0.2	74.3205	124.3421	0.7553	0.125
	100	0.5		124.3421		

Table 1 gives the means over 200 trials of the Wald statistics, of the p-value and the 5% empirical size of the Wald test. The column $\chi^2_{N,5\%}$ gives the critical values of the χ^2 distribution at the 5% level.

2. Estimation and tests on MFVB semiparametric random intercept and slope models for the application on neonates

In this section, we estimate Bayesian semiparametric random intercept and slope models (RCM) for infections and for accidents and we check if we accept the null hypothesis of Bayesian semiparametric random intercept and constant slope models (OWEC) for infections and for accidents. For the probability of MA, Table 2 gives the estimated coefficients[5], the standard deviations, the t-stats, the p-values and the 95% confidence intervals for β^R and β^G. Results are close to those of the semiparametric random intercept and slope model (RCM) for the probability of MA in the main text. The Wald test of the null hypothesis of an OWEC specification is $W = 6.1211$. As the degrees of freedom are $N(q^R - 1) = 3979 \times 11 = 43769$, then under the null, this statistic follows a $\chi^2_{N(q^R-1),5\%} = 44256.7943$ and the p-value is 0.00. We strongly accept the null hypothesis of a semiparametric random intercept and constant slope model for the occurrence of medical incidents/accidents (MA).

---

[5]To save space, we don't give any information on the estimated values of the $N(q-1)$ coeffcients u^R of the random part.

Table 2: Random intercept and slope model (RCM) for the probability of MA.

	Coef.	Sd.	t	P>|t|	min95%CI	max95%CI
intercept	-10.0033	0.3059	-32.7020	0.0000	-10.6028	-9.4037
prior infection	-0.5647	0.0385	-14.6740	0.0000	-0.6401	-0.4892
gestation age	0.0570	0.0071	8.0444	0.0000	0.0431	0.0709
newborn	-0.0727	0.0387	-1.8763	0.0606	-0.1486	0.0032
girl	0.0579	0.0249	2.3296	0.0198	0.0092	0.1066
cesarean	-0.0532	0.0313	-1.6997	0.0892	-0.1145	0.0081
twins	-0.0811	0.0321	-2.5244	0.0116	-0.1441	-0.0181
gravity	0.4833	0.0162	29.7712	0.0000	0.4514	0.5151
apgar	-0.2375	0.0317	-7.5010	0.0000	-0.2996	-0.1755
diagnostic related group class	0.3764	0.0324	11.6078	0.0000	0.3128	0.4399
occupation	-0.0224	0.0041	-5.4203	0.0000	-0.0305	-0.0143
admissions	-0.0084	0.0053	-1.5712	0.1161	-0.0188	0.0021
overtime hours	0.0060	0.0006	10.4510	0.0000	0.0049	0.0071
regular hours	0.0050	0.0003	19.0606	0.0000	0.0045	0.0055
weight	-0.0001	0.0000	-2.4126	0.0158	-0.0002	-0.0000
$\sigma_{u^R}^2 \times 10^3$	0.0050					
$\rho_{\sigma_{u^R}^2} \times 10^3$	2.2427					

Table 2 gives the estimated coefficients, the standard deviations, the t-stats, the p-values and the 95% confidence intervals for β^R and β^G. $\rho_{\sigma_{u^R}^2}$ is the proportion of the total variance contributed by the panel-level variance component $(\sigma_{u^R}^2)$: $\rho_{\sigma_{u^R}^2} = \sigma_{u^R}^2 / \left(\sigma_{u^R}^2 + \sigma_\varepsilon^2 \right)$ where $\sigma_\varepsilon^2 = \pi^2/3$ is the variance of the logistic distribution.

For the probability of HCAI, Table 3 gives the estimated coefficients, the standard deviations, the t-stats, the p-values and the 95% confidence intervals for β^R and β^G. Results are also close to those of the semiparametric random intercept and slope model (RCM) for the probability of HCAI in the main text except for `weight` and `gestation` whose coefficients are statistically not significantly different from zero. The Wald test of the null hypothesis of an OWEC specification is $W = 6.0227$. As the degrees of freedom are $N(q^R - 1) = 3979 \times 11 = 43769$, then under the null, this statistic follows a $\chi^2_{N(q^R-1),5\%} = 44256.7943$ and the p-value is 0.00. We also strongly accept the null hypothesis of a semiparametric random intercept and constant slope model for the occurrence of health care associated infections (HCAI).

Table 3: Random intercept and slope model (RCM) for the probability of HCAI.

	Coef.	Sd.	t	P>|t|	min95%CI	max95%CI
intercept	-7.3795	0.3342	-22.0786	0.0000	-8.0346	-6.7244
prior accident	-0.6357	0.0369	-17.2155	0.0000	-0.7081	-0.5634
gestation	-0.0209	223.6112	-0.0001	0.9999	-438.2907	438.2489
newborn	0.3254	0.0426	7.6473	0.0000	0.2420	0.4088
girl	-0.0180	0.0267	-0.6742	0.5002	-0.0704	0.0344
cesarean	-0.1925	0.0334	-5.7626	0.0000	-0.2579	-0.1270
twins	0.0542	0.0341	1.5891	0.1120	-0.0127	0.1211
occupation	0.0096	0.0041	2.3486	0.0188	0.0016	0.0176
admissions	0.0387	0.0056	6.9181	0.0000	0.0277	0.0496
apgar	-0.0165	0.0339	-0.4880	0.6256	-0.0830	0.0499
gravity	0.5860	0.0179	32.7750	0.0000	0.5510	0.6211
diagnostic related group class	0.3228	0.0359	8.9795	0.0000	0.2523	0.3932
overtime hours	0.0029	0.0006	4.8791	0.0000	0.0017	0.0041
regular hours	-0.0003	0.0003	-1.1227	0.2616	-0.0008	0.0002
weight	-0.0204	223.6112	-0.0001	0.9999	-438.2902	438.2494
$\sigma^2_{u^R} \times 10^3$	0.0082					
$\rho_{\sigma^2_{u^R}} \times 10^3$	2.8599					

Table 3 gives the estimated coefficients, the standard deviations, the t-stats, the p-values and the 95% confidence intervals for β^R and β^G. $\rho_{\sigma^2_{u^R}}$ is the proportion of the total variance contributed by the panel-level variance component ($\sigma^2_{u^R}$): $\rho_{\sigma^2_{u^R}} = \sigma^2_{u^R} / \left(\sigma^2_{u^R} + \sigma^2_{\varepsilon} \right)$ where $\sigma^2_{\varepsilon} = \pi^2/3$ is the variance of the logistic distribution.

3. Figures

1. Odds ratios for MA according to daily overtime hours and regular hours.
2. Odds ratios for MA according to daily overtime hours and birth weight.
3. Odds ratios for MA according to daily regular hours and birth weight.
4. Contours of odds ratios for MA according to regular and overtime hours (birth weight: decile 10).
5. Contours of odds ratios for MA according to regular and overtime hours (birth weight: decile 50).
6. Contours of odds ratios for MA according to regular and overtime hours (birth weight: decile 90).
7. Contours of odds ratios for MA according to overtime hours and birth weight (regular hours : decile 10).
8. Contours of odds ratios for MA according to overtime hours and birth weight (regular hours : decile 50).
9. Contours of odds ratios for MA according to overtime hours and birth weight (regular hours : decile 90).
10. Contours of odds ratios for MA according to regular hours and birth weight (overtime hours : decile 10).
11. Contours of odds ratios for MA according to regular hours and birth weight (overtime hours : decile 50).
12. Contours of odds ratios for MA according to regular hours and birth weight (overtime hours : decile 90).
13. Odds ratios for HCAI according to daily overtime hours and regular hours.
14. Odds ratios for HCAI according to daily overtime hours and birth weight.
15. Odds ratios for HCAI according to daily regular hours and birth weight.
16. Contours of odds ratios for HCAI according to regular and overtime hours (birth weight: decile 10).
17. Contours of odds ratios for HCAI according to regular and overtime hours (birth weight: decile 50).
18. Contours of odds ratios for HCAI according to regular and overtime hours (birth weight: decile 90).
19. Contours of odds ratios for HCAI according to overtime hours and birth weight (regular hours : decile 10).
20. Contours of odds ratios for HCAI according to overtime hours and birth weight (regular hours : decile 50).
21. Contours of odds ratios for HCAI according to overtime hours and birth weight (regular hours : decile 90).
22. Contours of odds ratios for HCAI according to regular hours and birth weight (overtime hours : decile 10).
23. Contours of odds ratios for HCAI according to regular hours and birth weight (overtime hours : decile 50).
24. Contours of odds ratios for HCAI according to regular hours and birth weight (overtime hours : decile 90).
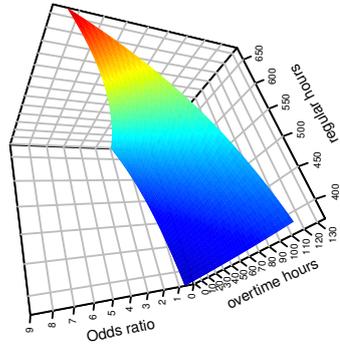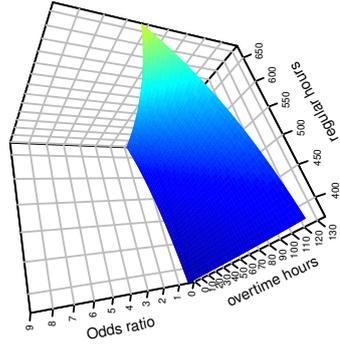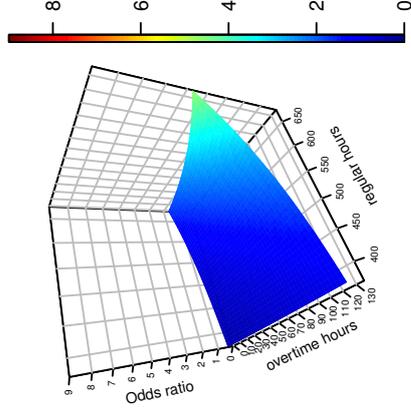
Figure A1: Odds ratios for MA according to daily overtime hours and regular hours.

Odds ratio for the probability of MA
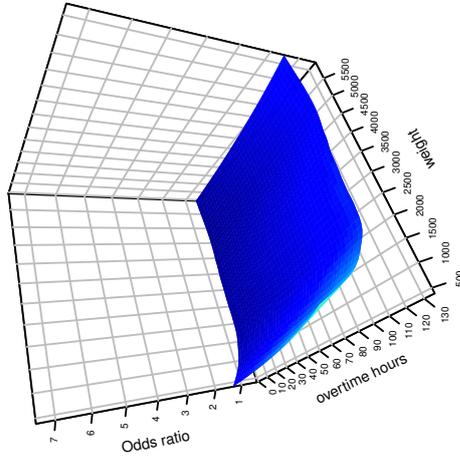
reference value of regular hours : decile 10 (455)
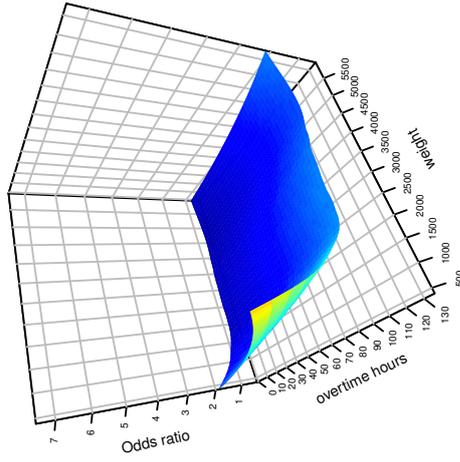
reference value of regular hours : decile 50 (523)
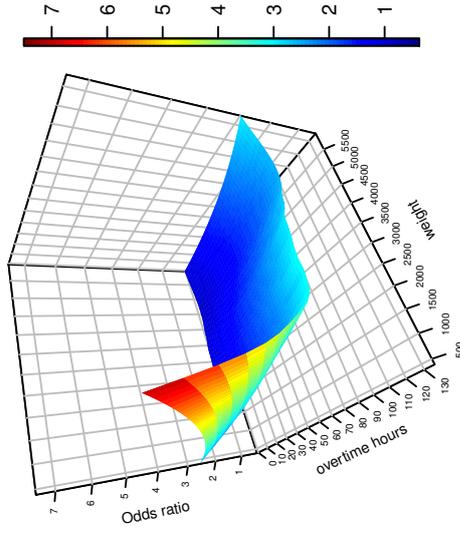
reference value of regular hours : decile 90 (584)

Figure A2: Odds ratios for MA according to daily overtime hours and birth weight.

10

Odds ratio for the probability of MA

reference value of overtime hours : decile 10 (2)

reference value of overtime hours : decile 50 (18)

reference value of overtime hours : decile 90 (53)

Figure A3: Odds ratios for MA according to daily regular hours and birth weight.
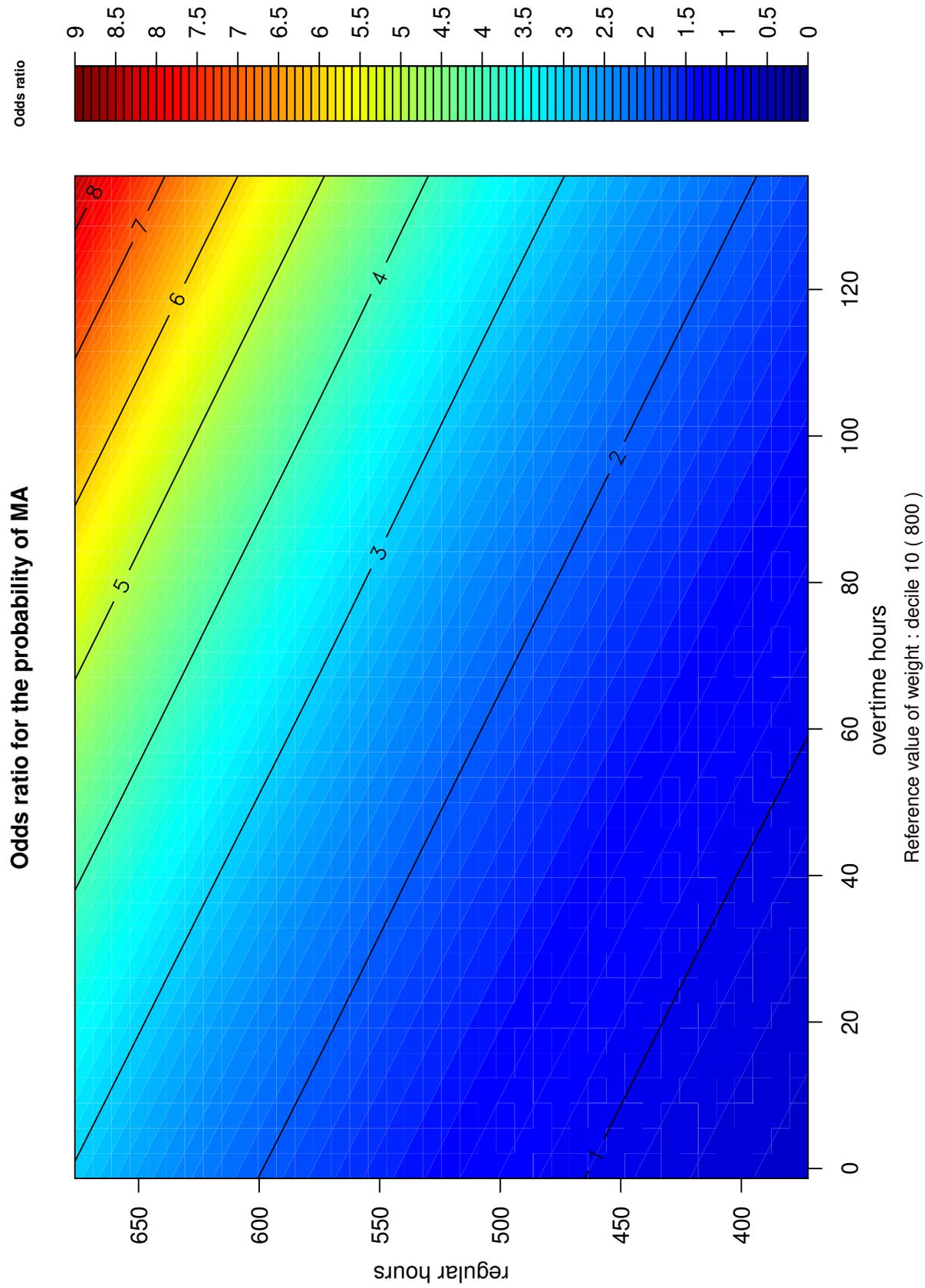
Figure A4: Contours of odds ratios for MA according to regular and overtime hours (birth weight: decile 10).
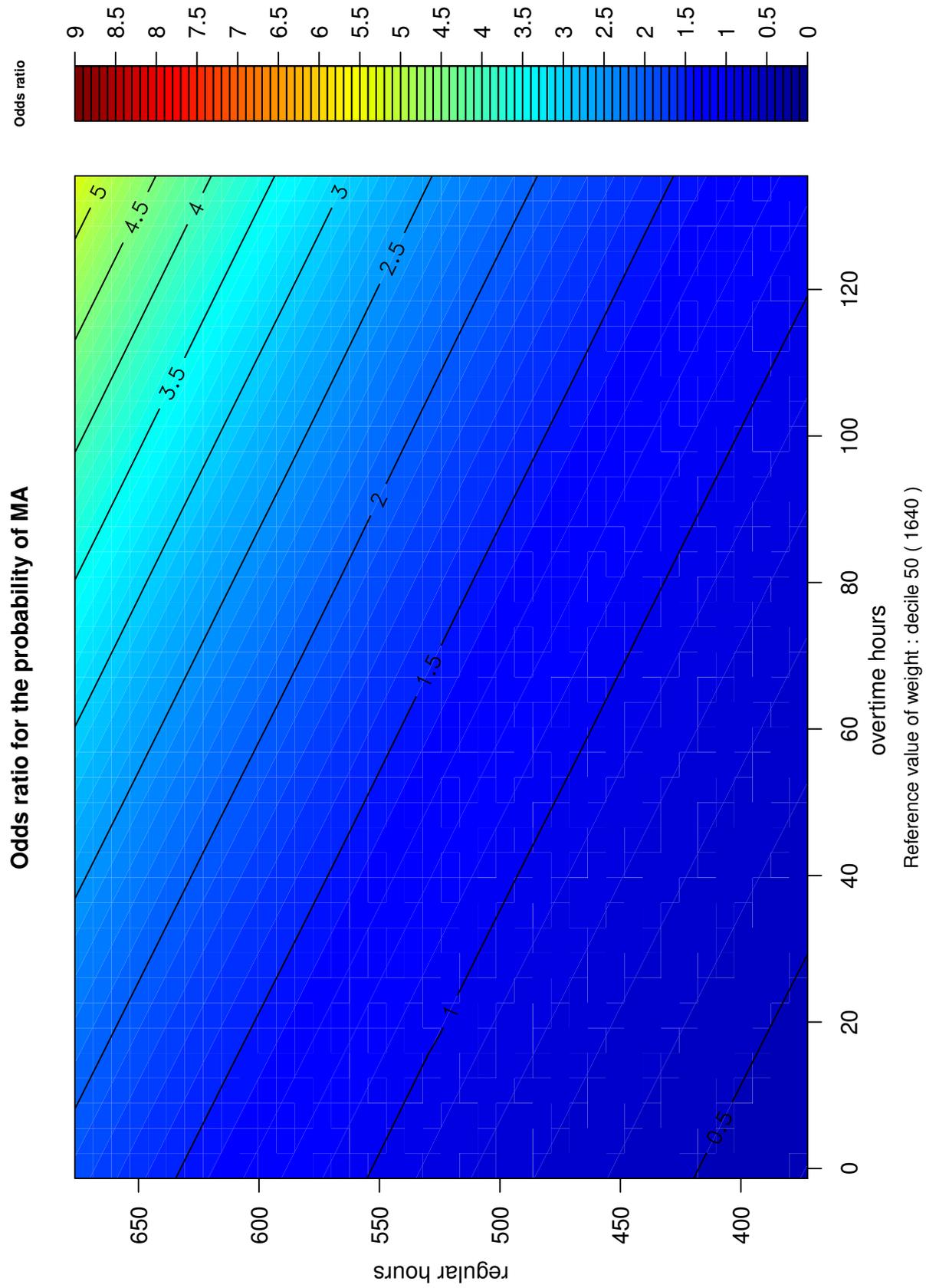
Figure A5: Contours of odds ratios for MA according to regular and overtime hours (birth weight: decile 50).
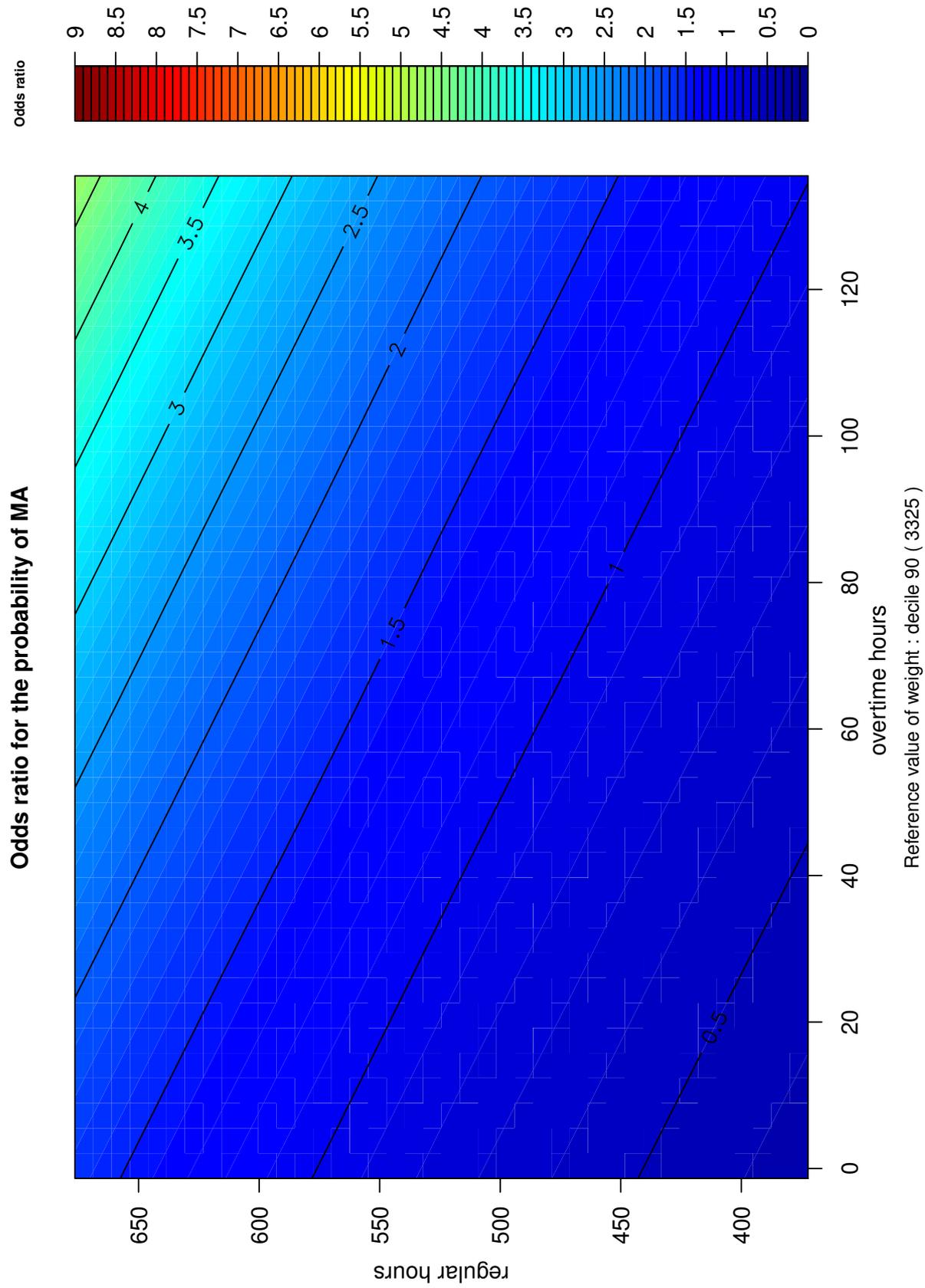
Figure A6: Contours of odds ratios for MA according to regular and overtime hours (birth weight: decile 90).
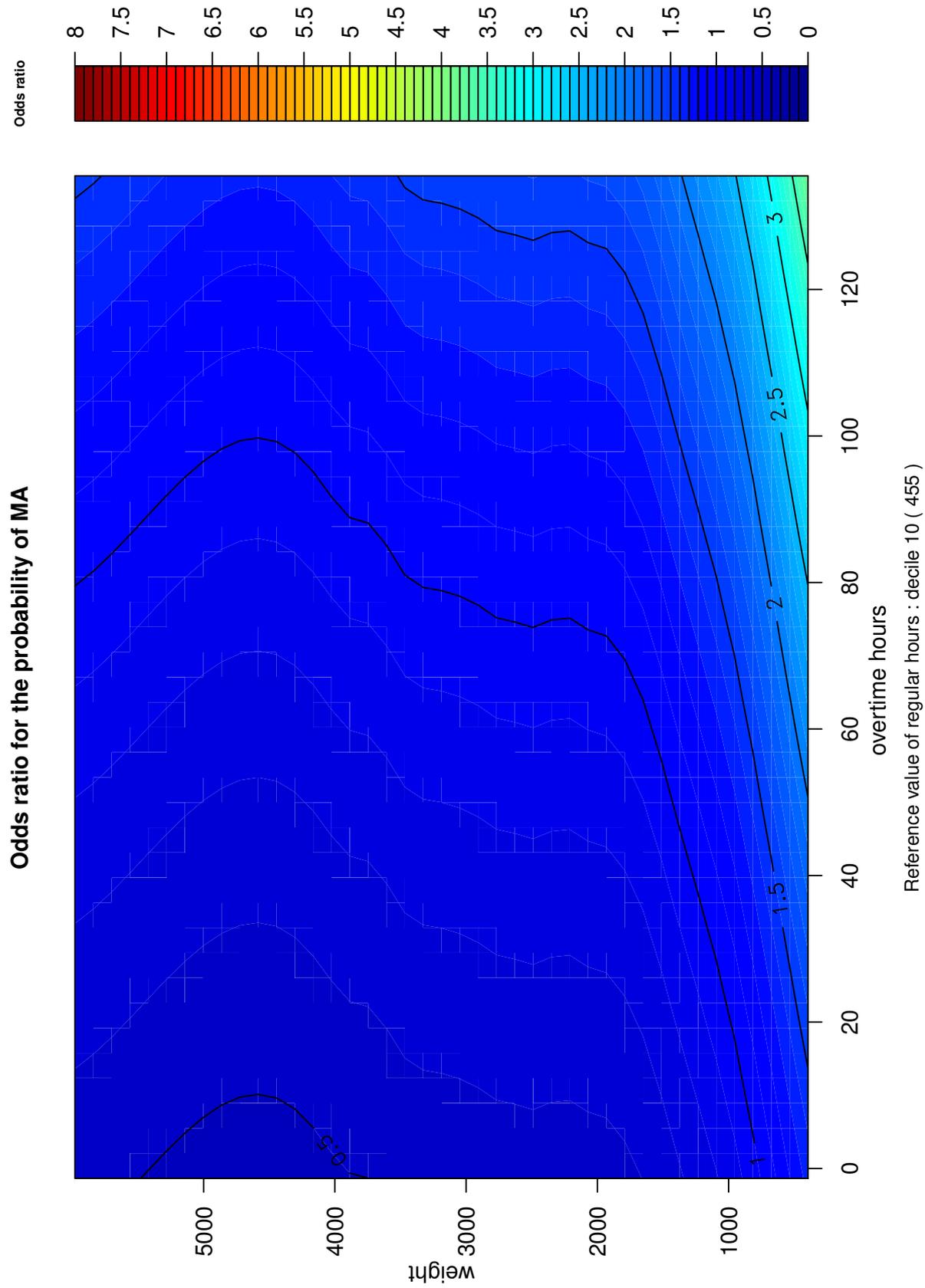
Figure A7: Contours of odds ratios for MA according to overtime hours and birth weight (regular hours : decile 10).
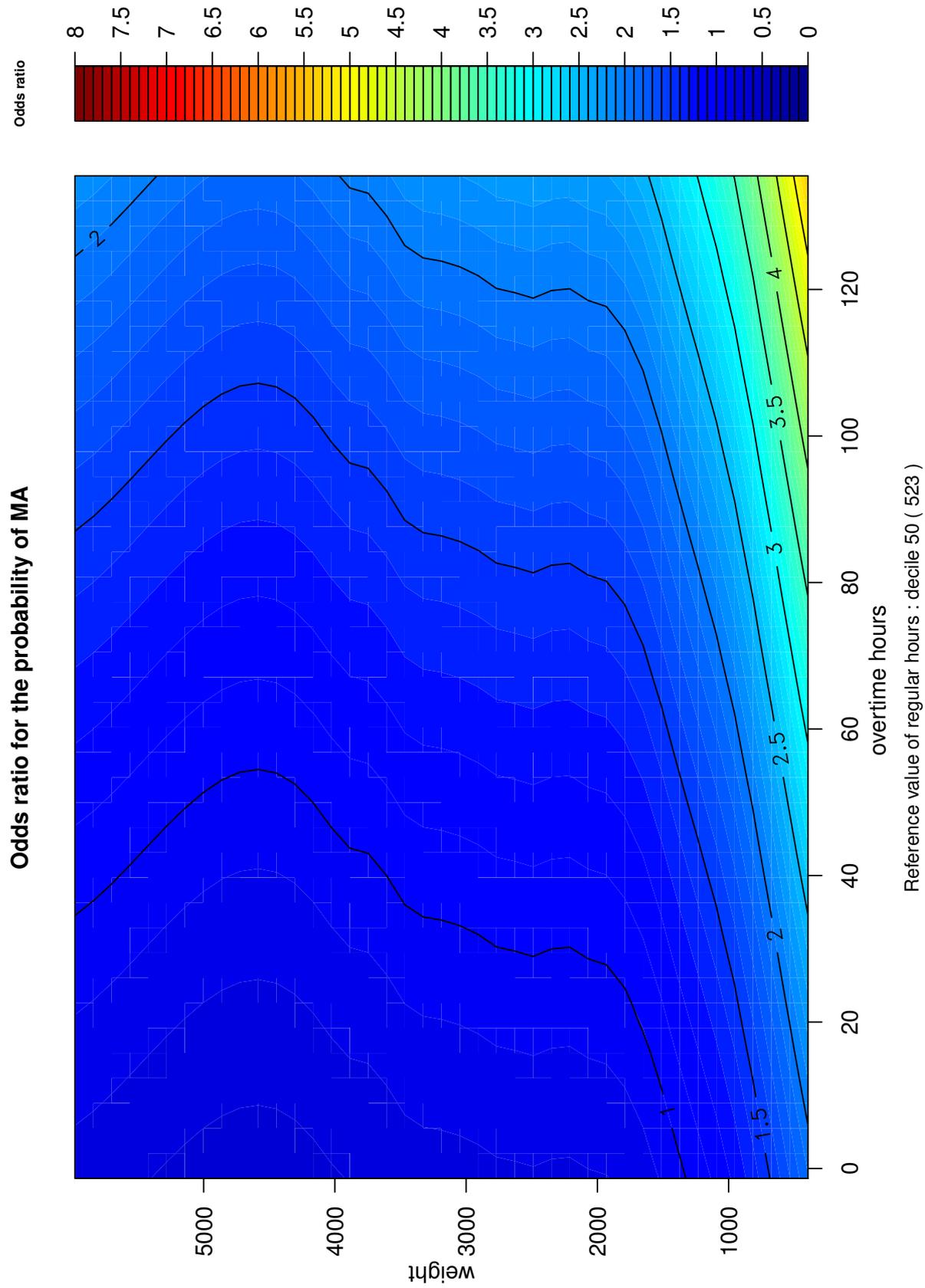
Figure A8: Contours of odds ratios for MA according to overtime hours and birth weight (regular hours : decile 50).
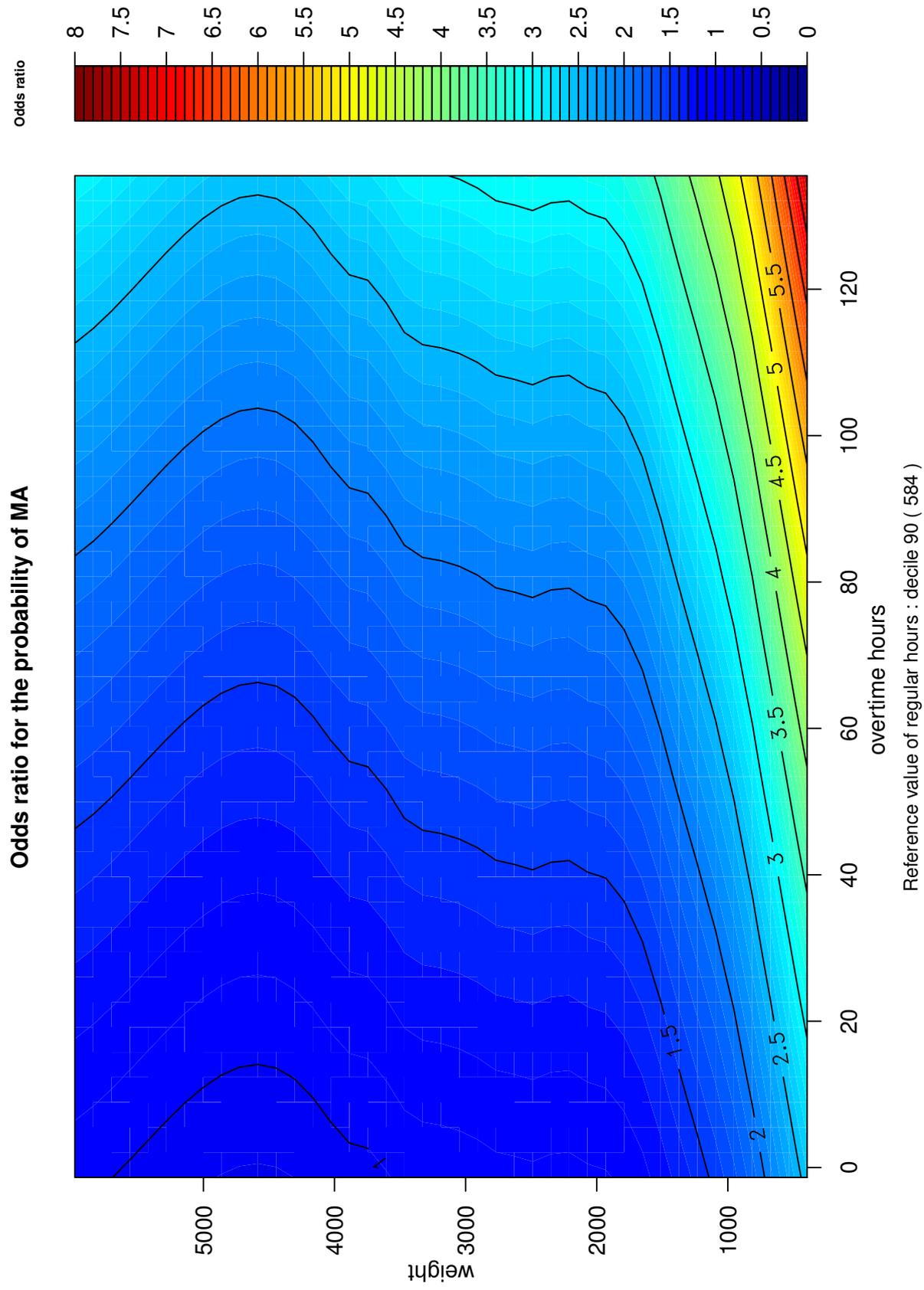
Figure A9: Contours of odds ratios for MA according to overtime hours and birth weight (regular hours : decile 90).
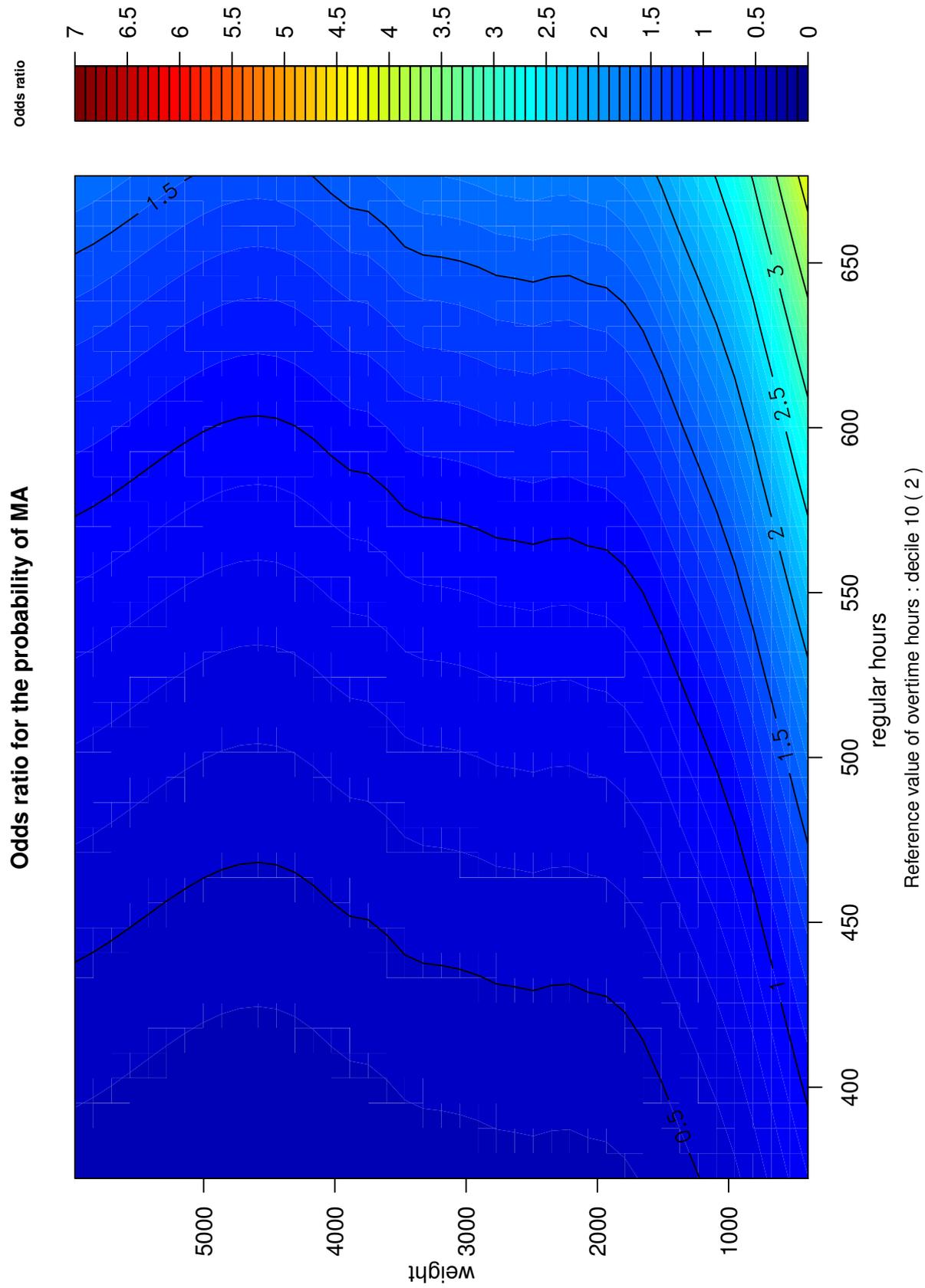
Figure A10: Contours of odds ratios for MA according to regular hours and birth weight (overtime hours : decile 10).

Figure A11: Contours of odds ratios for MA according to regular hours and birth weight (overtime hours : decile 50).
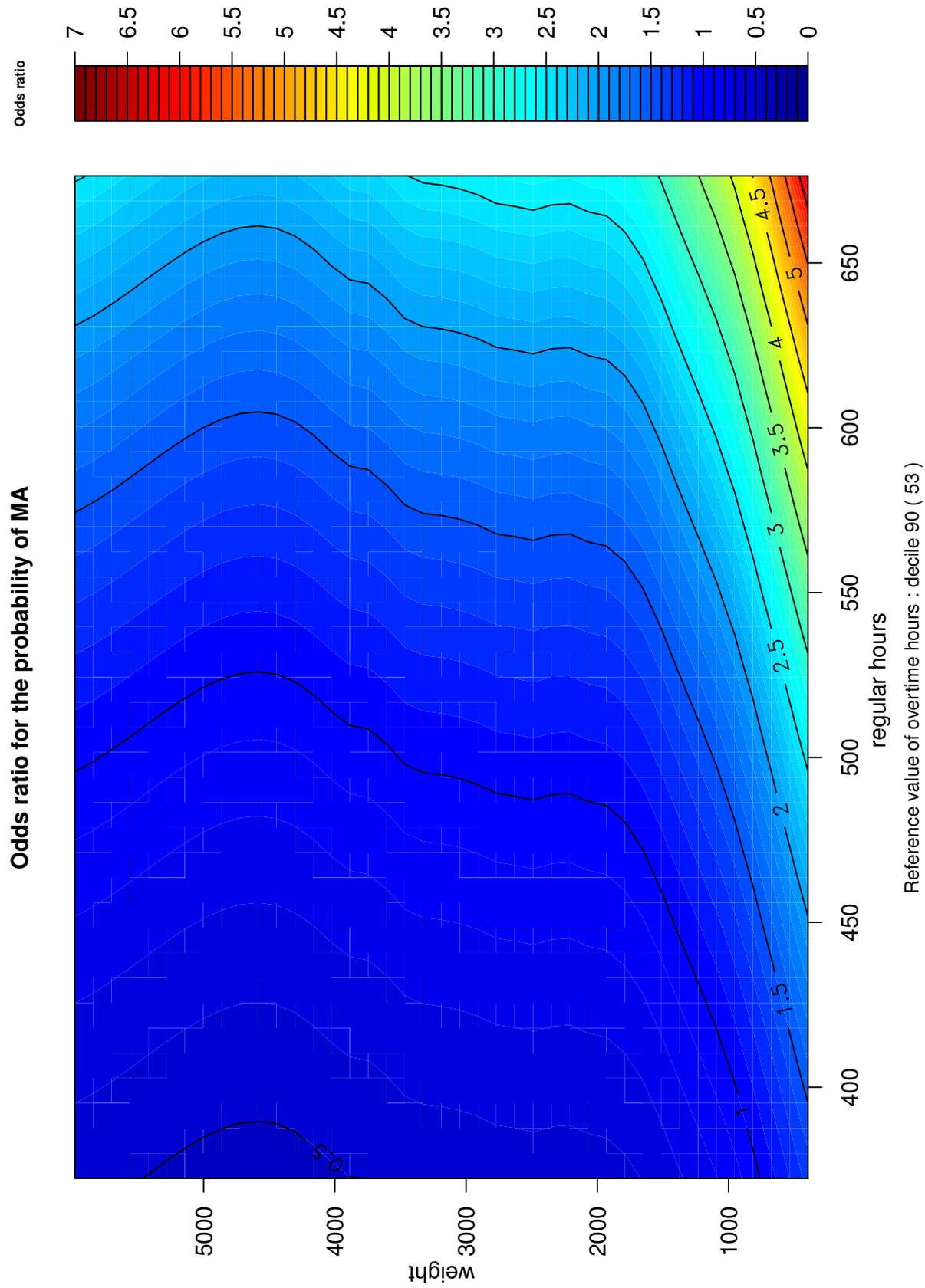
Figure A12: Contours of odds ratios for MA according to regular hours and birth weight (overtime hours : decile 90).
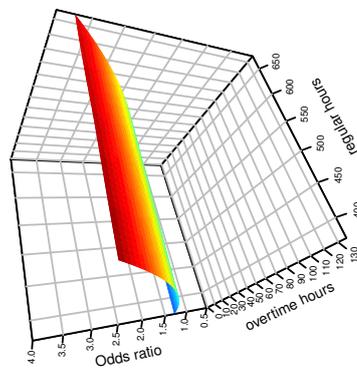
Figure A13: Odds ratios for HCAI according to daily overtime hours and regular hours.

Figure A14: Odds ratios for HCAI according to daily overtime hours and birth weight.

Odds ratio for the probability of HCAI
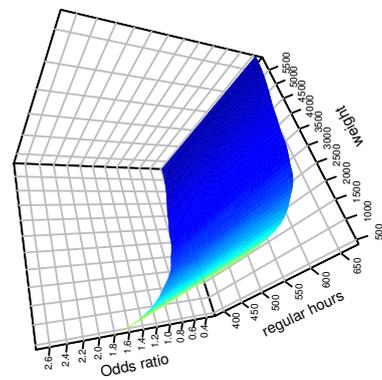
reference value of overtime hours : decile 10 (2)
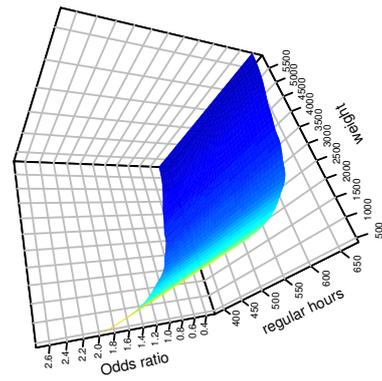
reference value of overtime hours : decile 50 (18)

reference value of overtime hours : decile 90 (53)

Figure A15: Odds ratios for HCAI according to daily regular hours and birth weight.
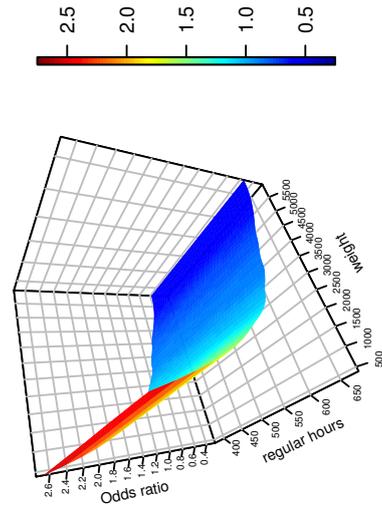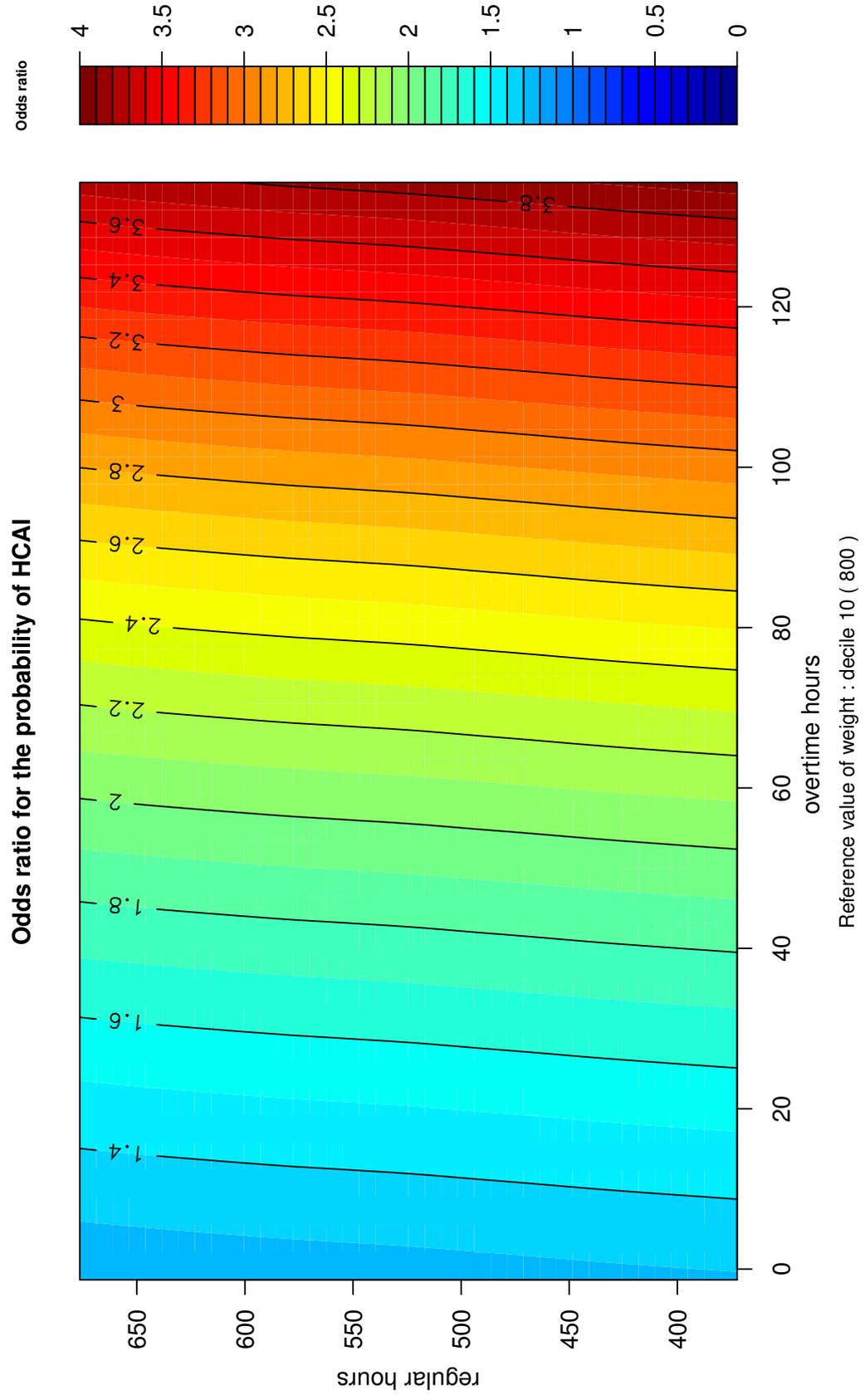
Figure A16: Contours of odds ratios for HCAI according to regular and overtime hours (birth weight: decile 10).
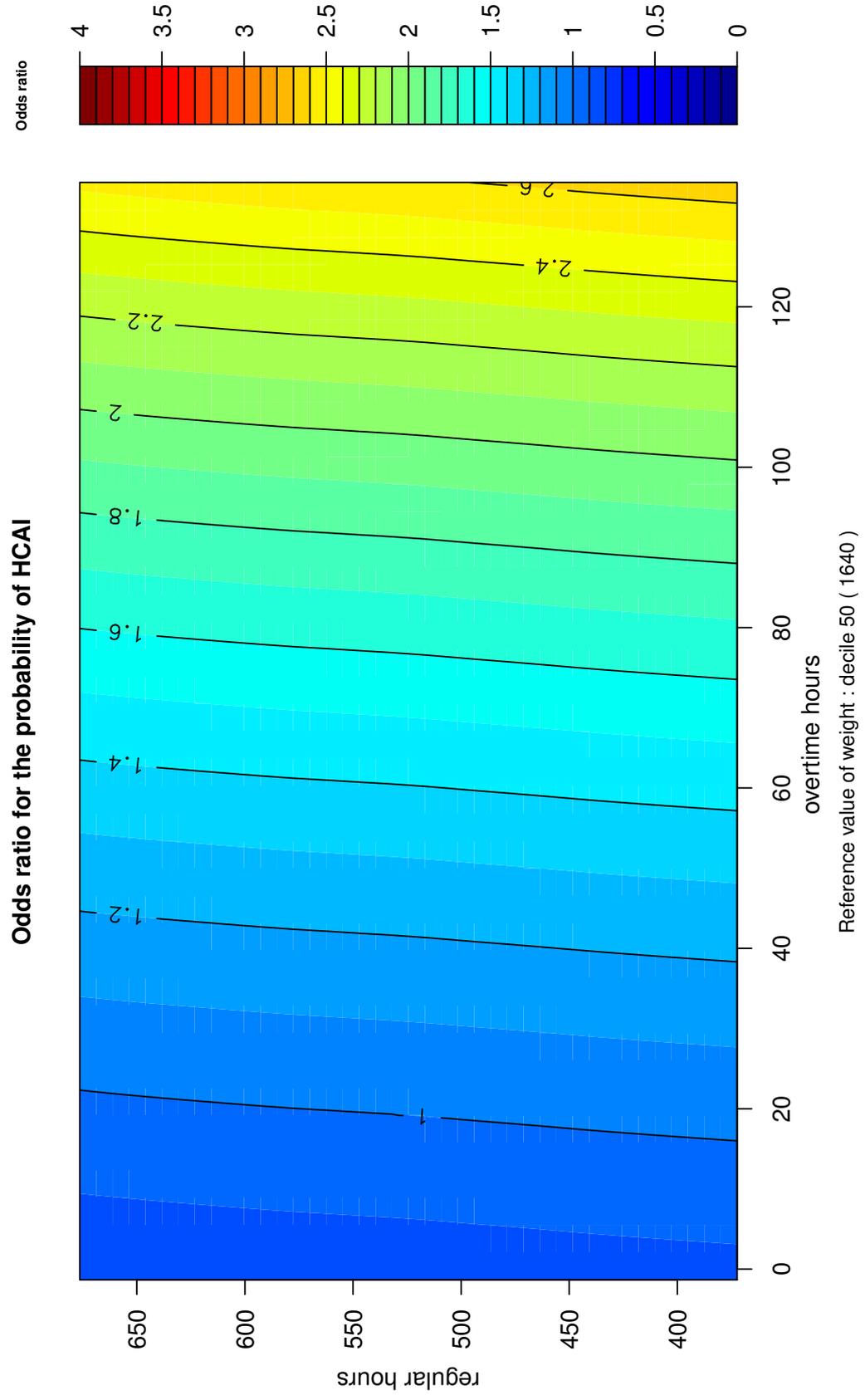
Figure A17: Contours of odds ratios for HCAI according to regular and overtime hours (birth weight: decile 50).
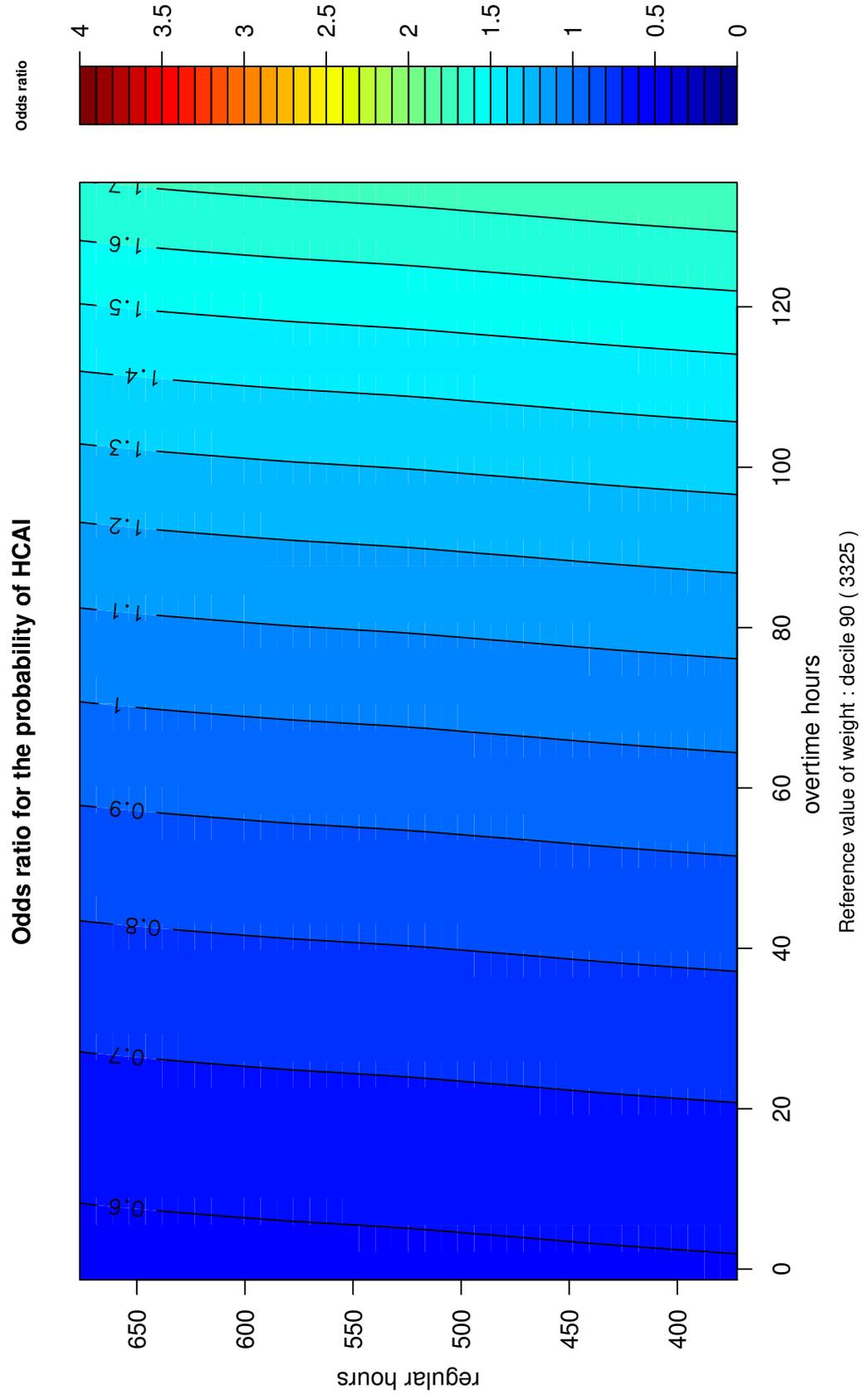
Figure A18: Contours of odds ratios for HCAI according to regular and overtime hours (birth weight: decile 90).
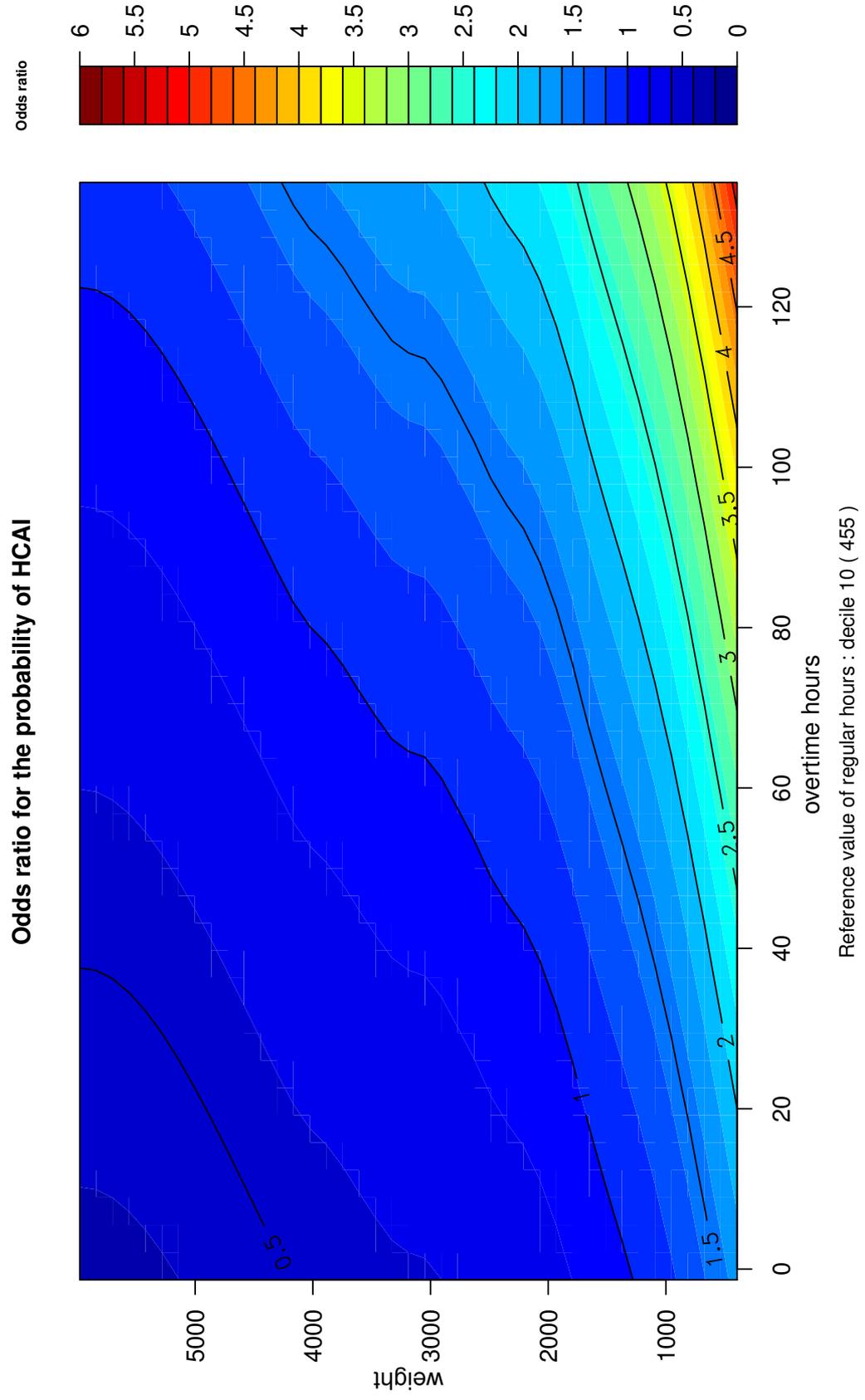
Figure A19: Contours of odds ratios for HCAI according to overtime hours and birth weight (regular hours : decile 10).
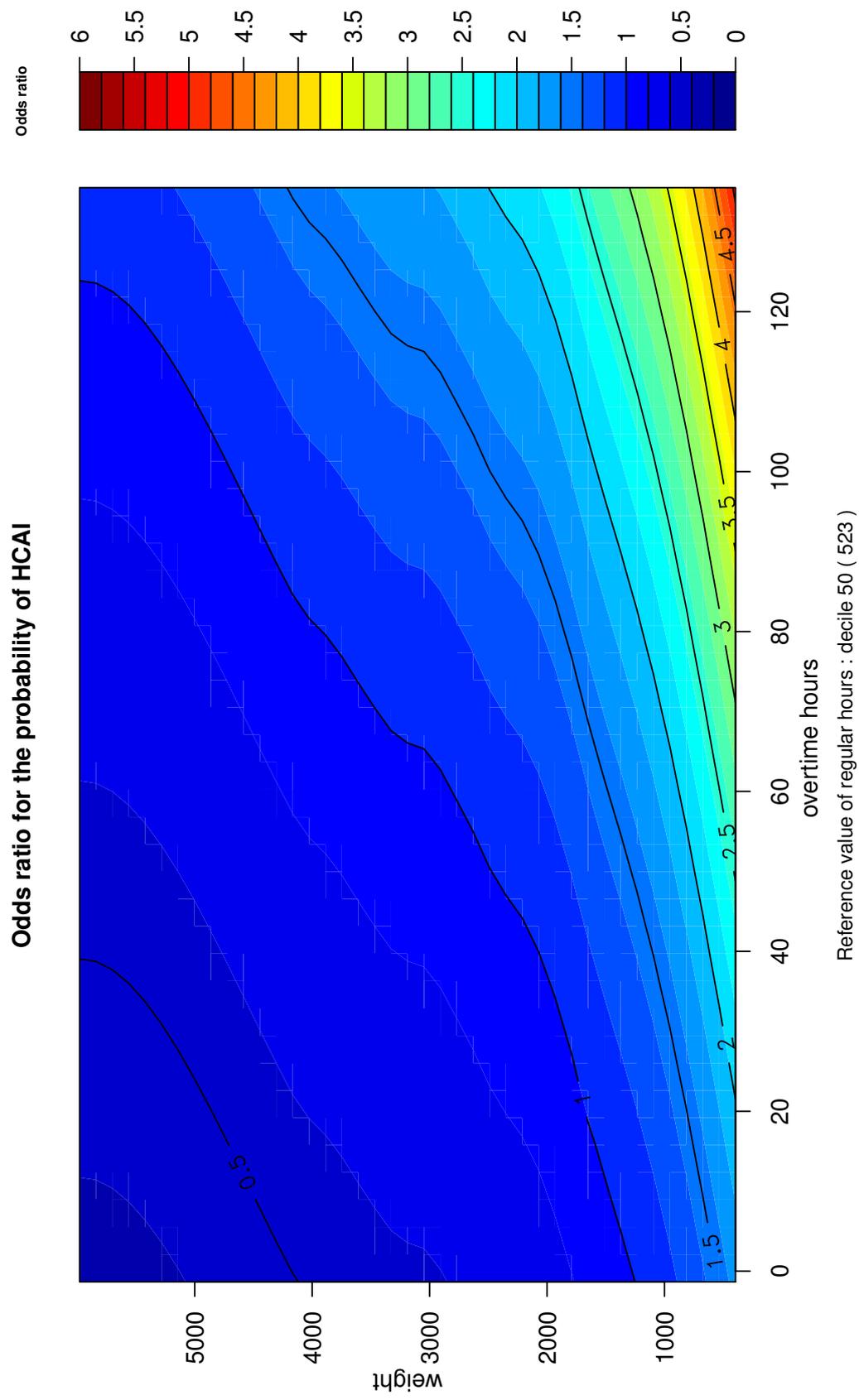
Figure A20: Contours of odds ratios for HCAI according to overtime hours and birth weight (regular hours : decile 50).
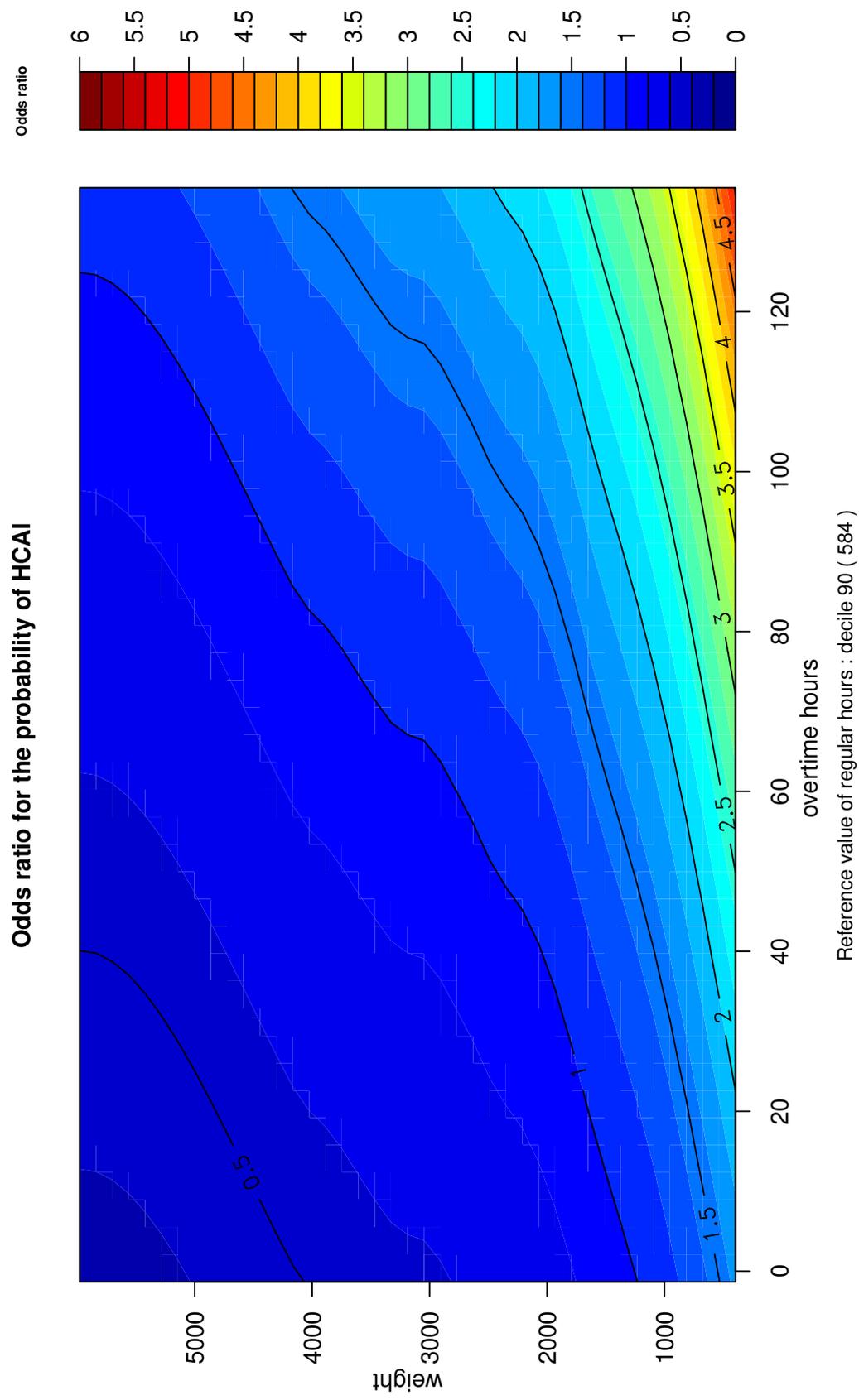
Figure A21: Contours of odds ratios for HCAI according to overtime hours and birth weight (regular hours : decile 90).
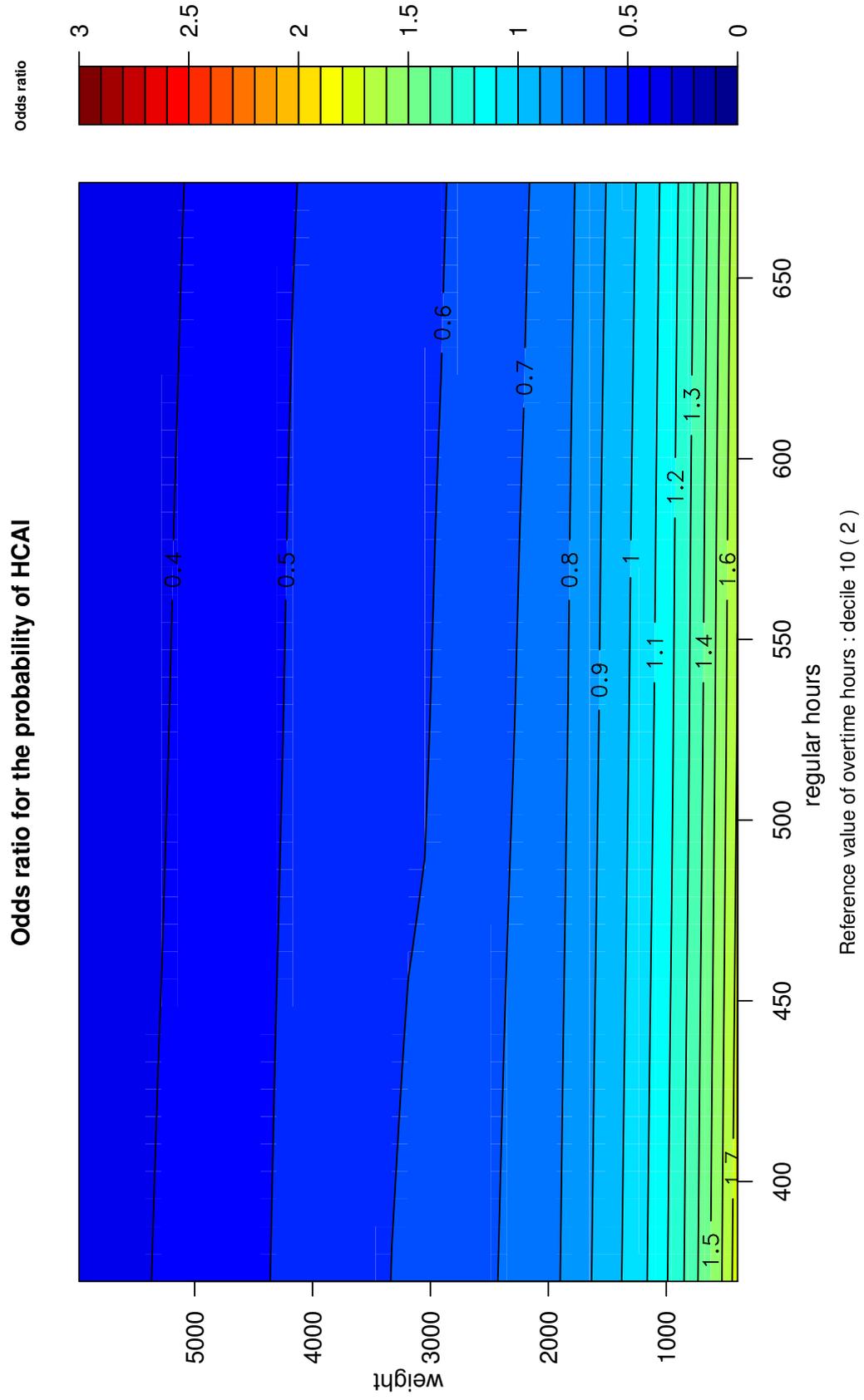
Figure A22: Contours of odds ratios for HCAI according to regular hours and birth weight (overtime hours : decile 10).
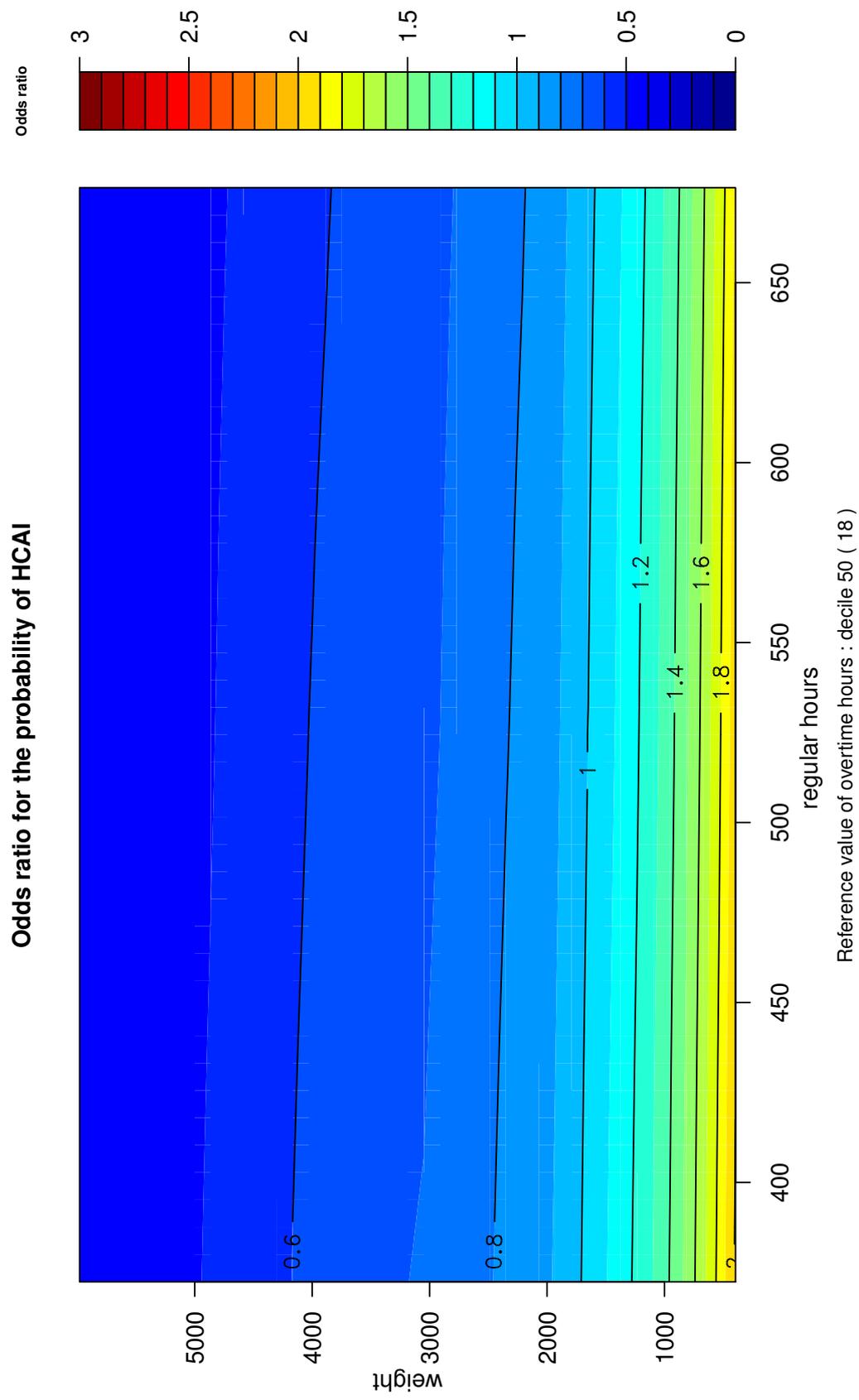
Figure A23: Contours of odds ratios for HCAI according to regular hours and birth weight (overtime hours : decile 50).
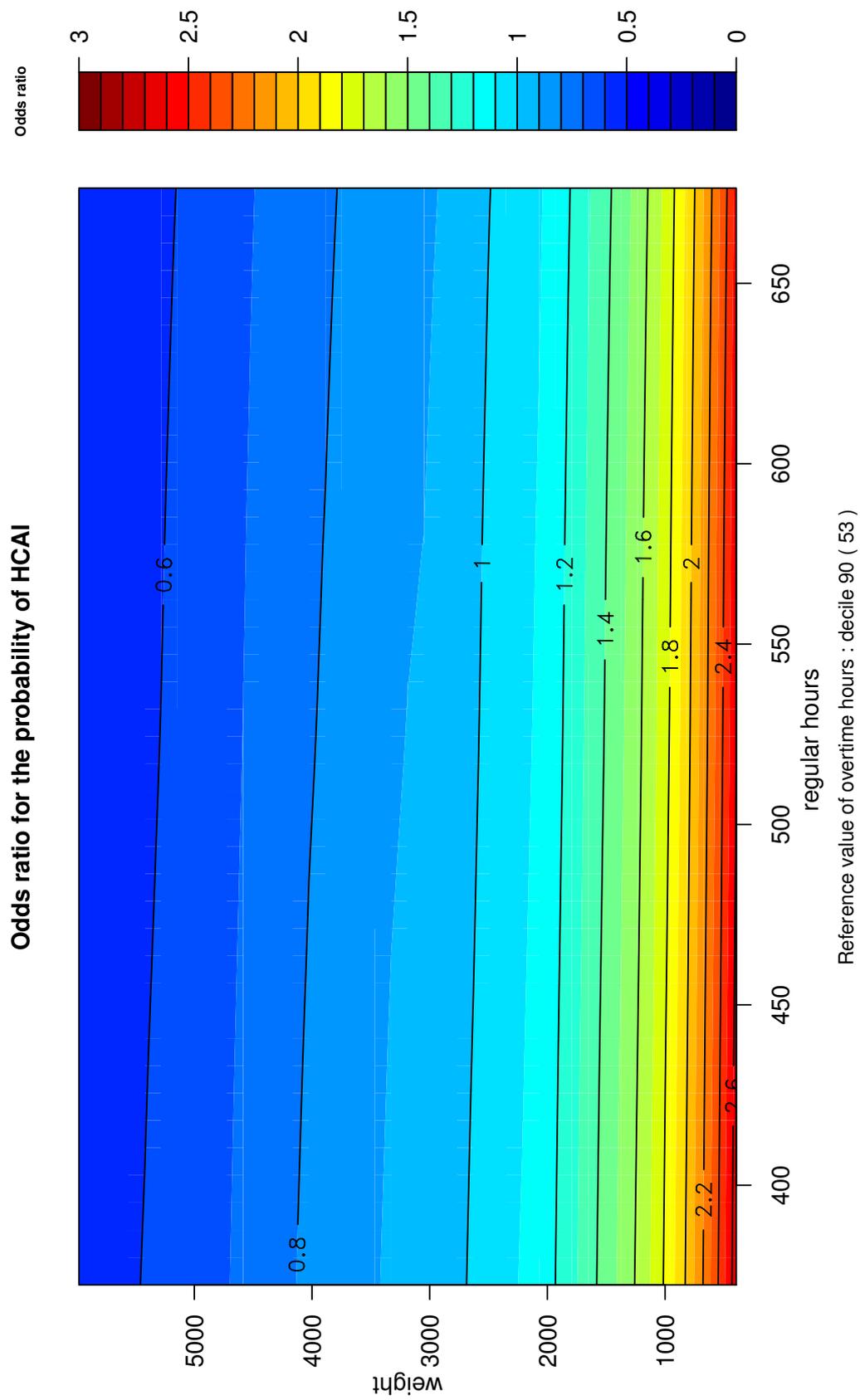
Figure A24: Contours of odds ratios for HCAI according to regular hours and birth weight (overtime hours : decile 90).

4. References

Beal, M., Ghahramani, Z., 2003. The variational Bayesian EM algorithm for incomplete data: with application to scoring graphical model structures, in: Bernardo, J., Bayarri, M., Berger, J., Dawid, A., Heckerman, D., Smith, A., West, M. (Eds.), Bayesian Statistics. Oxford University Press, Oxford, UK. volume 7, pp. 453–464.

Chérief-Abdellatif, B.E., Alquier, P., 2018. Consistency of variational Bayes inference for estimation and model selection in mixtures. Electronic Journal of Statistics 12, 2995–3035.

Dieng, A.B., Tran, D., Ranganath, R., Paisley, J., Blei, D., 2017. Variational inference via χ upper bound minimization, in: Guyon, I., Luxburg, U.V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R. (Eds.), Advances in Neural Information Processing Systems 30. Curran Associates, Inc., pp. 2732–2741.

Gregory, A., Veall, M., 1985. Formulating Wald tests of nonlinear restrictions. Econometrica 53, 1465–1468.

Li, Z., Sillanpää, M., 2013. A Bayesian nonparametric approach for mapping dynamic quantitative traits. Genetics 194, 997–1016.

Nott, D.J., Tran, M.N., Leng, C., 2012. Variational approximation for heteroscedastic linear models and matching pursuit algorithms. Statistics and Computing 22, 497–512.

Yao, Y., Vehtari, A., Simpson, D., Gelman, A., 2018. Yes, but did it work?: Evaluating variational inference. arXiv preprint arXiv:1802.02538 .