2011

Bicycle Commuting in Melbourne During the 2000s Energy Crisis: A Semiparametric Analysis of Intraday Volumes

Michael S Smith, *Melbourne Business School*
Goeran Kauermann

Bicycle Commuting in Melbourne During the 2000s Energy Crisis: A Semiparametric Analysis of Intraday Volumes

Michael S. Smith[a,⋆] and Göran Kauermann[b]

[a] Melbourne Business School, University of Melbourne

[b] Department of Business Administration and Economics, University of Bielefeld

8th July 2011

To Appear in the Journal of Transportation Research Part B: Methodological

⋆ Corresponding Author; address for correspondence: Professor Michael Smith, Melbourne Business School, University of Melbourne, 200 Leicester Street, Carlton, VIC 3053, Australia. Email: mike.smith@mbs.edu

Abstract

Cycling is attracting renewed attention as a mode of transport in western urban environments, yet the determinants of usage are poorly understood. In this paper we investigate some of these using intraday bicycle volumes collected via induction loops located at ten bike paths in the city of Melbourne, Australia, between December 2005 and June 2008. The data are hourly counts at each location, with temporal and spatial disaggregation allowing for the impact of meteorology to be measured accurately for the first time. Moreover, during this period petrol prices varied dramatically and the data also provide a unique opportunity to assess the cross-price elasticity of demand for cycling. Over-dispersed Poisson regression models are used to model volumes at each location and at each hour of the day. Seasonality and the impact of weather conditions are modelled as semiparametric and estimated using recently developed multivariate penalized spline methodology. Unlike previous studies that use aggregate data, the empirical results show a substantial meteorological and seasonal component to usage. They also suggest there was substitution into cycling as a mode of transport in response to increases in petrol prices, particularly during peak commuting periods and by commuters originating in wealthy and inner city neighbourhoods. Last, we extend the approach to a multivariate longitudinal count data model using a Gaussian copula estimated by Bayesian data augmentation. We find first order serial dependence in the hourly volumes and a 'return trip' effect in daily bicycle commutes.

Key Words: Cross-Price Elasticity; Discrete Copula Model; Generalized Mixed Models; Mode of Transport; Multivariate Penalized Spline Smoothing; Count Data Model.

1 Introduction

Cycling offers many advantages as a mode of transport in western society. These include zero marginal emissions of greenhouse gases and pollutants, reduced traffic congestion and numerous health benefits, including reduced rates of obesity; see Pucher, Dill & Handy (2010) for extensive evidence on the health benefits of cycling. All of these are major contemporary issues, and it has lead to an increase in interest in the determinants of bicycle usage; for example, see Rietveld & Daniel (2004), Wardman, Tight & Page (2007) Hunt & Abraham (2007) and Heinen et al. (2010). To date, studies in the transportation literature have almost exclusively focused on the analysis of survey or experimental data which allow for the study of individual level factors on the decision to cycle; for example, see Nankervis (1999), Stinson & Bhat (2004), Shannon et al. (2006) and Xing et al. (2010). Analysis of data on cycling volumes that are aggregate with respect to individuals, but disaggregate over time and location, has not been attempted for two reasons. The first is a lack of appropriate data, while the second is the difficulty in controlling for the complex impact of weather conditions. This paper addresses both problems by exploiting a new and unique source of data, and by employing multivariate penalized spline smoothing to account for the impact of meteorology and seasonality.

We analyse hourly cycling volumes at ten locations on dedicated bicycle paths in the city of Melbourne in Australia. These locations have been selected to provide an accurate picture of inner-city bicyle commutes and are collected automatically using recently installed induction loops placed under each path; in this sense the data are both comprehensive and highly accurate.[1] We examine hourly cyclist counts from December 2005 to June 2008, which corresponds to a period in which oil prices experienced extreme levels of volatility and finishes just before the peak of US$147.29 per barrel on 11 July 2008.[2] As a result, during this period

[1] These counters have been installed by VicRoads and they believe these count the number of cyclists who pass over them with over 95% accuracy.

[2] This is the all time peak at the time of writing.

average Melbourne unleaded petrol pump prices varied between a minimum of 108.5 cents per litre to 161.0 cents per litre. Therefore, this data provides a unique opportunity to assess the cross-elasticity of the demand for cycling with respect to petrol prices. We model hourly counts using over-dispersed Poisson regression models. Each hour of the day is modelled separately, thereby controlling for the strong diurnal variation in all aspects of the model. Such a separation is also common in the intraday modelling of demand for flow commodities, such as electricity (Fiebig, Bartels & Aigner 1991; Cottet & Smith 2003). The mean of each Poisson model contains linear effects for trend, day type and the logarithm of a measure of Melbourne petrol prices, so that the corresponding cross-elasticity is constant. The effect of seasonality and intraday weather conditions are accounted for as semiparametric effects.

While transport researchers suspect weather and seasonality affect the propensity to cycle in quite different ways than other modes of transport (Heinen et al. 2010), to date there is only limited evidence to this effect. For example, Nankervis (1999) and Shannon et al. (2006) find minor and no weather effects, respectively, using limited datasets on Australian university students. In our study, by matching hourly observations from local weather stations with those from the induction loops we find that both immediate weather conditions and season are major determinants of bicycle commutes. In each Poisson model the impact of weather is modelled as an unknown multivariate function of six meteorological variables. We follow an approach advocated in the environmental (Shively & Sager 1999) and econometric (Panagiotelis & Smith 2008) modelling literatures and simplify the effect as additive in univariate main effects and all pairwise bivariate interactions. This results in 21 unknown univariate and bivariate functions which are estimated jointly using penalized spline smoothing. Here, each function is modelled using a high-dimensional basis representation with unknown coefficients that are penalized with a quadratic penalty to ensure a smooth fit. Such an approach has a long history in the statistical literature, see Wahba (1978), Wecker & Ansley (1983) and O'Sullivan (1988) for early examples, but has seen much recent attention and extension; see the overviews by Ruppert, Wand & Carroll (2003; 2009) and Wood (2006).

We follow Wahba (1978), Wong & Kohn (1996), Wood (2003), Wand (2003) and Lang & Brezger (2004) and interpret the penalties as Gaussian prior distributions for the basis coefficients, so that the Poisson model can be interpreted as a generalized linear mixed model (GLMM). Scalar 'smoothing parameters' control the level of smoothing of each function, and these can be estimated from the data using a marginalized penalized likelihood derived from a Laplace approximation to the full penalized likelihood. The use of the Laplace approximation in such contexts is justified asymptotically in Kauermann, Krivobokova & Fahrmeir (2009) and Rue, Martino & Chopin (2009). The use of approximations to a full likelihood function is increasingly popular when the full likelihood function is computationally intractable. For example, in the transportation literature Ferdous et al. (2010) approximate the likelihood of a multivariate ordered probit model as a composition of bivariate ordered probit model likelihoods, which is easily maximized to provide a consistent estimator. Following Wager, Vaida & Kauermann (2007) we show how the marginalized likelihood can be used to compute a marginal Akaike Information Criteria (AIC). Using this criteria we adopt a stepwise algorithm to identify spurious components in the additive decomposition of the meteorological effect. The end result is a flexible, smooth and parsimonious representation of the multivariate effect of meteorology on cycling volumes that can be estimated joint with other aspects of the Poisson model. Only by carefully controling for the complex nonlinear impact of weather and season can reliable estimates of the cross-elasticities of demand with respect to petrol prices be obtained.

We find that both over-dispersion and intraday serial dependence are present in the hourly volumes. To account for the latter we construct a 14-dimensional copula model for the longitudinal vector of hourly volumes during the day. Copula models are multivariate models that can be constructed from arbitrary marginal models, with any dependence captured by a copula function. Copula modeling is increasingly popular throughout the physical and social sciences, and has been employed previously in transport studies by Bhat & Eluru (2009) and Sener et al. (2010). In our study the margins are the over-dispersed Poisson regression

models, while a Gaussian copula function is used to account for the intraday dependence. The Gaussian copula function is used here because it allows for both positive and negative dependencies between counts at all hours, and can be used in higher dimensions unlike many alternatives. The end result is a multivariate longitudinal Poisson model for dependent counts, where the semiparametric exogenous effects have forms that vary over each hour. As highlighted by Danaher & Smith (2011) it is not widely appreciated that many existing latent variable based count models can be viewed as special cases of a copula model. For example, Song (2000) expresses the multivariate probit model as a Gaussian copula. Similarly, the correlated count model of Herriges et al. (2008) is an example of a Gaussian copula model.

As discussed in Danaher & Smith (2011) because the margins are discrete-valued, estimation of the dependence parameters of the copula using direct maximum likelihood is computationally impractical for a 14-dimensional Gaussian copula. We therefore follow Pitt, Chan & Kohn (2006) and Danaher & Smith (2011) and estimate the Gaussian copula using data augmentation and Bayesian Markov chain Monte Carlo (MCMC). The approach of estimating the marginal models first, and then the copula parameters in a second step, is widely employed in the copula literature; see Cherubini, Luciano & Vecchiato (2004) and Silva & Lopez (2008) for discussions. Estimates of pairwise Spearman correlations reveal strong first order serial dependence between the hourly volumes, along with additional dependence between volumes during the morning and evening commuting peaks. The latter is a 'return trip' effect evident in the counts, but caused at the individual level where people who commute by bicycle to work are more likely to return by the same mode of transport than would otherwise be the case. To show the flexibility of the copula model we compute the resulting distribution of total daily volumes for any given day, and also the forecasts of evening volumes given those observed in the morning.

The paper continues by outlining the data, and then the over-dispersed Poisson model. Multivariate penalized splines are discussed, along with their interpretation as a generalized linear mixed model. We show how to estimate the smoothing parameters using the marginal-

ized likelihood, which is derived in greater detail in Appendix A. Using the marginalized likelihood we also define the marginalized AIC criteria. The intraday analysis of the effects of weather, season and petrol prices on bicycle commutes follow, along with an outline of the copula model and the rich intraday serial dependence structure uncovered. Appendix B provides a summary of the Bayesian MCMC scheme used to estimate the Gaussian copula models, while a discussion concludes.

2 Data

2.1 Bicycle Usage in Melbourne

The cycling data we examine were obtained from VicRoads, which is a branch of the state government of Victoria. As well as responsibility for the management of the state road network, the organisation oversees other transport initiatives.[3] These include the development and implementation of cycling programs throughout Victoria, with the aim of increasing participation in cycling as part of an integrated transport network (Austroads Inc., 2005). Melbourne is the capital of Victoria, with a population of 3.81 million people in 2007, which makes up 73.1% of the state total.[4] The city is located at the mouth of the Yarra river with a flat geography which is well-suited to cycling, and possesses an extensive bicycle network. Nevertheless, cycling is still only a minor mode of transport in Melbourne, comprising only 1.5% of total trips to work in 2006 (Bonham & Suh 2008), and the city has a high level of car dependency and a growing traffic congestion problem (Clarke & Hawkins 2005).

From 2005 VicRoads began to install inductive loop counters on off-road bicycle paths at key locations surrounding the Melbourne central business district (CBD). These count bicycle volumes with a high degree of accuracy at an intraday level. We examine hourly counts of cyclists at the ten locations where loops were first installed and where the data is most complete. Table 1 lists these locations, along with the orientation of the location

[3]For further information visit their extensive website at www.vicroads.vic.gov.au.

[4]These are Australian Bureau of Statistics estimates at 30 June 2007; ABS, 'Population by Age and Sex, Regions of Australia' (ABS cat. no. 3235.0).

in relation to the CBD and the direct line distance between the loop and the General Post Office (GPO) building located in the centre of the CBD. The number of bicycle trips is low both between 19:00 – 06:00 and during weekends, and we focus here only on weekday trips between 06:00 – 19:00 on working days, which are dominated by daily commutes. Table 1 lists mean hourly counts during these hours, and also broken down by the morning peak period 06:00 – 09:00 and evening peak period 16:00 – 19:00. Also listed are the number of observations and time frame over which the data are available.[5]

—Table 1 about here—

2.2 Spatial Variation

Melbourne has a radial transport network (Clarke & Hawkins 2005), and the loops have been located to count commuting trips to and from the CBD and immediate locale. There is a strong segmentation of the city suburbs into four quadrants radiating from the CBD, which roughly correspond to the four directions of the compass. This division is based upon topology, patterns of urban development and local government boundaries, and is associated with substantial differences in the socio-economic status of residents. The southern and eastern suburbs are more affluent than the northern and western suburbs, with average incomes of residents up to 50.5% higher (ABS 2005).[6]

Each loop is located on an off-road cycle path which radiates between the CBD and from suburbs in one of the four quadrants, as listed in Table 1. Using this information, we create several aggregate counts based upon spatial location. The first two are based upon the direction of origin (the south and east suburbs, SE, and the northern suburbs, N), while a third is for inner city loops (IC) closest to the CBD. In creating these counts, we exclude

[5]We exclude the annual 'Ride to Work Day' and the days during the 2008 Melbourne Commonwealth Games.

[6]The 2005 average total income of wage earners for the Melbourne ABS statistical regions are: Western Melb. $42,358; Moreland City (Inner North) $40,907; Northern Middle Melb. $42,985; Boroondara City (Inner East) $61,555; Eastern Middle Melb. $45,731; Southern Melb. $51,695.

the three loops with the highest proportion of missing data (numbers 1, 4 and 8), as well as omitting days where any of the component loops in the cumulative counts were missing. Last, we also create a city-wide count (*ALL*) at the seven loops with the most complete data. Summaries for these four cumulative counts are also provided in Table 1.

2.3 Seasonality and Time of Day

Figure 1 contains boxplots of the hourly city-wide counts (*ALL*) broken down by both season and time of day. The four panels correspond to the southern hemisphere seasons of Spring (September to November), Summer (December to February), Autumn (March to May) and Winter (June to August). Seasonal variation is apparent, with lower volumes being observed in winter compared to summer. Counts vary strongly according to the time of day, and peak during the morning (07:00 – 09:00) and evening (16:00 – 19:00) commuting periods.

——Figures 1 and 2 about here.——

2.4 Meteorology and Cost of Fuel

Even a cursory examination of this data reveal that meteorological conditions are a key determinant of bicycle commutes. For example, Figure 2 plots the hourly counts at the St Georges Road loop on the 21 November in both 2006 and 2007. Both days were working weekdays, yet counts in 2007 were substantially lower than in 2006. This is likely due to meteorological variation, with it being cool and raining between 07:00 and 19:00 on this day in 2007, yet hot and dry on 2006. The relationship between weather conditions and the propensity to cycle is indiscernible in previous studies that use either individual level survey data (Shannon et al. 2006; Hunt & Abraham 2007), or data that is highly aggregated with respect to time or location (Nankervis 1999; Rietveld & Daniel 2004).

Hourly data was obtained on meteorological variables from the Australian Bureau of Meteorology observed at the three Melbourne metropolitan weather stations located closest to the

loops.[7] The variables considered were temperature ($TEMP$), humidity (HMD), windspeed ($WIND$) and rainfall ($RAIN$). To account for a small number of missing observations, and minor spatial variation, we employ average measurements across the three stations. The manner in which meteorological conditions affect the decision to cycle to work (rather than switch to other modes of transport) is complex, and to help capture this we introduce two additional variables. The first is a measure of recent rainfall ($RRAIN$), which is the total rainfall in the current and two previous hourly periods, while the second is the daily maximum temperature ($MTEMP$).

The petrol price data we examine are average weekly pump prices of unleaded petrol observed at retail outlets in the Melbourne metropolitan region.[8] Unleaded petrol is by far the largest category of fuel sold in Australia, and has a price that is almost perfectly correlated with that of other fuels, such as diesel, LPG and higher octane unleaded petrol. Figure 3 plots the average real weekly petrol price in July 2008 Australian dollars[9] ($PETROL$) and the total weekday counts at all loops. The linear correlation between real petrol prices and total daily counts is statistically insignificant, with any relationship being hard to determine because of the very high variance in counts. Last, we note here that while the automobile is by far the most popular mode of transport in Melbourne, the city also has an extensive integrated public transport system with a single ticketing system. However, the real price of public transport usage remained constant during the period [10] and we therefore do not include this in our model.

——Figure 3 about here——

[7]The stations were Essendon Airport, Viewbank and Moorabbin Airport. Melbourne CBD was not used because its measurements are known to be unreliable due to the surrounding high rise buildings.

[8]This data are constructed by the Australian monitoring organization FuelTrac; see www.fueltrac.com.au.

[9]The price deflator used was the Australian Bureau of Statistics quarterly consumer price index for all groups.

[10]Prices were increased inline with inflation three times during the period: on 1 January 2006, 3 June 2007 and 1 January 2008.

3 Modelling Usage

3.1 Poisson Regression Model

We model counts separately at each hour of the day for each loop, or cumulative count. Denoting y_i as the ith observation of an hourly count, for $i = 1, \ldots, n$ days, we employ an over-dispersed Poisson regression model with moments

$$E(y_i|\eta_i) = \mu_i = \exp(\eta_i)\,,\ \mathrm{Var}(y_i|\eta_i) = \mu_i\phi\,, \tag{3.1}$$

and $\phi \geq 1$ an unknown dispersion parameter (Wedderburn 1974). Estimation is undertaken using the Fisher scores (see McCullagh & Nelder 1989; p.40)

$$\tilde{y}_i = \eta_i + (y_i - \mu_i)/\mu_i\,, \quad \text{for } i = 1, \ldots, n\,, \tag{3.2}$$

where $E(\tilde{y}_i) = \eta_i$ and $\mathrm{Var}(\tilde{y}_i) = \phi/\mu_i$. Following many previous authors we compute inference using quasi-maximum likelihood (QMLE) assuming normality for $\tilde{y}_i$, so that

$$\tilde{y}_i \sim N(\eta_i, \phi/\mu_i). \tag{3.3}$$

Fisher scoring proceeds iteratively by (i) calculating $\tilde{y} = \{\tilde{y}_1, \ldots, \tilde{y}_n\}$ and (ii) fitting $\eta = \{\eta_1, \ldots, \eta_n\}$ using the Gaussian likelihood at equation (3.3) for given $\tilde{y}$; see Wolfinger & O'Connell (1993) for a discussion.

We model the mean as

$$\eta_i = z_i'\alpha + s(t_i) + m(x_i)\,, \tag{3.4}$$

where z_i is a vector of linear effects with coefficients α, $s(t_i)$ is a smooth seasonal effect with respect to the time of year $0 \leq t_i < 1$, and $m(x_i)$ is an unknown multivariate function of the six meteorological variables $x_i = (x_{i1}, \ldots, x_{i6})$. This is an extension of the generalized

linear model (Wood 2006; Chapter 3) to include flexible semiparametric components s and m. In the linear component we include an intercept, linear time trend, five day type dummy variables for Tuesday to Friday and for school holidays. We also include $\log(PETROL)$, so that the cross-price elasticity of the mean

$$\frac{\partial \mu_i}{\partial PETROL_i} \times \frac{PETROL_i}{\mu_i} = \alpha_P$$

is constant, resulting in a total of 8 linear terms. While it is possible to also model the time trend as semiparametric, we do not do so because it would only be identifiable from the seasonal component by the smoothness parameter of s. Similarly, the elasticity can also be modelled as semiparametric, but we do not do so here because it is both a weak effect, and is harder to interpret if nonlinear.

The effect of weather upon counts is complex and highly nonlinear. To simplify the problem we assume additivity up to bivariate interactions between meteorological variables, so that

$$m(x_i) = \sum_{j=1}^{6} m_j(x_{ij}) + \sum_{j=1}^{6} \sum_{l>j} m_{jl}(x_{ij}, x_{il}) \,, \tag{3.5}$$

where m_j and m_{jl} are univariate and bivariate smooth functions, respectively. Similar additivity assumptions have been used to model the effect of meteorological variables upon ozone levels (Shively & Sager 1999) and intraday electricity demand (Panagiotelis & Smith 2008). The additive decomposition at equation (3.5) involves the estimation of 21 unknown functions and we identify a more parsimonious representation in a data-based fashion. Let $\mathcal{I} = \{1, 2, \ldots, 6\}$ be the index set of the meteorological variables. We denote with $\mathcal{S} \subset \mathcal{I}$ the set of indices where m_j are nonlinear functions, and with $\mathcal{B} = \{(i, j); i \neq j, i \in \mathcal{S}, j \in \mathcal{S}\}$ the set of non-zero bivariate functions. Further denoting the model $\mathcal{M} = \{\mathcal{S}, \mathcal{B}\}$, we define

a parsimonious representation of equation (3.5) as

$$m(x_i; \mathcal{M}) = \sum_{j \in \mathcal{I}} x_{ij}\beta_j + \sum_{j \in \mathcal{S}} m_j(x_{ij}) + \sum_{(j,l) \in \mathcal{B}} m_{jl}(x_{ij}, x_{il}). \tag{3.6}$$

3.2 Penalized Splines

We discuss determination of $\mathcal{M}$ in Section 3.4, but outline here how the unknown semi-parametric components in equation (3.6) can each be expressed as penalized splines for given $\mathcal{M}$. We represent each unknown function as a linear combination of thin plate spline basis terms (Wahba 1990, pp.30–34), which also corresponds to the popular cubic smoothing spline basis for the univariate functions. We also use the same basis for the seasonal effect s in equation (3.4), but with periodicity enforced on the basis terms so that $s(0) = s(1)$. The functions are therefore

$$s(t_i) = \tilde{b}_s(t_i)'\tilde{v}_s\,,\ m_j(x_{ij}) = \tilde{b}_j(x_{ij})'\tilde{v}_j\,,\ \text{ and } m_{jl}(x_{ij}, x_{il}) = \tilde{b}_{jl}(x_{ij}, x_{il})'\tilde{v}_{jl}\,, \tag{3.7}$$

for $j \in \mathcal{S}$ and $(j, l) \in \mathcal{B}$. The vectors $\tilde{b}_s, \tilde{b}_j, \tilde{b}_{jl}$ are bases evaluated at the ith observation, and $\tilde{v}_s, \tilde{v}_j, \tilde{v}_{jl}$ are coefficient column vectors for which we employ quadratic penalty terms. For the univariate functions, this penalty is equivalent to the integrated squared second order derivative of the function (Wahba 1978; Green & Silverman 1994, p.13). For the bivariate functions we penalize the sum of the integrated squared elements of the Hessian matrix (Wood 2003).

The knots of the thin plate basis terms are located at all unique observations of the independent variables, so that there can be up to n terms in the basis. To reduce the computational burden we follow Hastie (1996) and Wood (2003) and employ low rank smoothing. For each function, this involves a spectral decomposition of the design matrix for the basis decomposition. The k eigenvectors corresponding to the k largest eigenvalues are used as a reduced

rank basis, so that we redefine the functions in equation (3.7) as

$$s(t_i) = b_s(t_i)'v_s\,,\ m_j(x_{ij}) = b_j(x_{ij})'v_j\,,\ \text{and}\ m_{jl}(x_{ij}, x_{il}) = x_{ij}x_{il}\beta_{jl} + b_{jl}(x_{ij}, x_{il})'v_{jl}\,. \quad (3.8)$$

Here, β_{jl} is an unpenalized linear interaction coefficient and v_s, v_j, v_{jl} are basis coefficient k-vectors for the reduced rank basis terms b_s, b_j, b_{jl}. The reduced rank basis coefficients have quadratic penalties $v_s'D_sv_s$, $v_j'D_jv_j$ and $v_{jl}'D_{jl}v_{jl}$, where D_s, D_j and D_{jl} are diagonal matrices containing the k largest eigenvalues for the spectral decomposition of each basis. Such an approach is similar to that advocated by Eilers & Marx (1996), Ruppert, Wand & Carroll (2003), Lang & Brezger (2004) and others who directly select k basis terms and an associated quadratic penalty with non-diagonal matrices. We follow Wood (2006, p.161) and set $k = 10$ for the univariate functions and $k = 30$ for the bivariate functions, and there is substantial evidence that selecting k larger has little influence on the fit; see Ruppert (2002) or Kauermann & Opsomer (2011) for discussions.

Let $\lambda_s, \lambda_j, \lambda_{jl}$ be smoothing parameters for functions s, m_j, m_{jl}, respectively. For model $\mathcal{M}$, let $\theta_\mathcal{M}$ denote both the basis term coefficients and linear terms for all components in equations (3.4), (3.6) and (3.8), and $\lambda_\mathcal{M}$ be the set of all smoothing parameters. Then the penalized quasi-log-likelihood takes the form

$$l_p(\theta_\mathcal{M}, \lambda_\mathcal{M}) = l(\theta_\mathcal{M}) - \frac{1}{2}\sum_{j\in\mathcal{S}}\lambda_j v_j'D_jv_j - \frac{1}{2}\sum_{(j,l)\in\mathcal{B}}\lambda_{jl}v_{jl}'D_{jl}v_{jl} - \frac{1}{2}\lambda_s v_s'D_sv_s\,, \quad (3.9)$$

where $l(\theta_\mathcal{M}) = \sum_{i=1}^n l_i(\theta_\mathcal{M})$ is the unpenalized quasi-log-likelihood from equation (3.3) for a given model $\mathcal{M}$. Conditional upon the smoothing parameters $\lambda_\mathcal{M}$, maximization of equation (3.9) with respect to $\theta_\mathcal{M}$ can be undertaken via Fisher scoring as discussed in Section 3.1. The estimation of $\lambda_\mathcal{M}$ and selection of $\mathcal{M}$ is discussed below.

3.3 Generalized Linear Mixed Model

Following Wahba (1978), Wong & Kohn (1996), Wood (2003) and others we interpret the penalties in equation (3.9) as Gaussian priors for the reduced rank basis term coefficients, so that

$$v_s \sim N(0, \lambda_s^{-1} D_s^{-1}), \quad v_j \sim N(0, \lambda_j^{-1} D_j^{-1}) \quad \text{and} \quad v_{jl} \sim N(0, \lambda_{jl}^{-1} D_{jl}^{-1}), \tag{3.10}$$

thereby reformulating the problem as a GLMM. Estimation proceeds by maximizing the logarithm of the marginal likelihood $l_m(\alpha, \beta_{\mathcal{M}}, \lambda_{\mathcal{M}}; \mathcal{M})$ obtained by integrating out the reduced rank basis coefficients. This integration can be undertaken based on the Laplace approximation as suggested by Breslow & Clayton (1993) and outlined in more detail in Appendix A. Asymptotic results supporting this approach in the penalized spline smoothing context can be found in Kauermann, Krivobokova & Fahrmeir (2009) and Rue, Martino & Chopin (2009).

Maximization of the Laplace approximation of l_m with respect to $\{\alpha, \beta_{\mathcal{M}}, \lambda_{\mathcal{M}}\}$ is pursued as part (ii) of the Fisher scoring algorithm outlined in Section 3.1 using the standard steps for a linear mixed model (LMM) listed in Harville (1977); also see Appendix A for more details. Re-calculation of $\tilde{y}_1, \ldots, \tilde{y}_n$ and refitting the LMM provide the essential steps of the Fisher scoring algorithm. The smoothing parameters $\lambda_{\mathcal{M}}$ are reciprocals of prior variances which are estimated following optimisation. The procedure is implemented numerically in regular statistical software, where we make use of the 'gamm' procedure provided in the package 'mgcv' in the statistical language 'R'.[11] Finally, the semiparametric function estimates can be computed via penalized least squares estimation of the basis term coefficients in equation (3.9). The benefits for writing the model in the mixed model framework are twofold. First, we directly obtain estimates for the smoothing parameters $\lambda_{\mathcal{M}}$, thereby enabling computation of the function estimates. Second, we can also use the mixed model framework for model selection as discussed below.

[11] Implementation of GLMMs and smoothing in R is discussed by Wood (2006).

3.4 Model Selection

The index set $\mathcal{M} = \{\mathcal{S}, \mathcal{B}\}$ determines the functional form of the meteorological effect m in equation (3.6), and we determine this from the data. We use the AIC (Akaike 1974) computed using the marginal likelihood; a criterion that is labelled marginal AIC by Vaida & Blanchard (2005) and Wager, Vaida & Kauermann (2007), and has been recently extended to GLMMs by Lavergne, Martinez & Trottier (2008). For model $\mathcal{M}$, let $l_m(\hat{\alpha}, \hat{\beta}_{\mathcal{M}}, \hat{\lambda}_{\mathcal{M}}; \mathcal{M})$ be the maximum of the logarithm of the (Laplace approximation) to the marginal likelihood at step (ii) of the Fisher scoring algorithm. Then we define the marginal AIC as

$$mAIC(\mathcal{M}) = -2l_m\big(\hat{\alpha}, \hat{\beta}_{\mathcal{M}}, \hat{\lambda}_{\mathcal{M}}; \mathcal{M}\big) + 2q\,, \tag{3.11}$$

where q is the number of components (linear or nonlinear) in the model at equation (3.6). We could calculate *mAIC* for all possible models $\mathcal{M}$, but we constrain the search strategy as follows. First, for numerical stability we do not consider models with bivariate effects in the pairs (*RAIN*,*RRAIN*) and (*TEMP*,*MTEMP*) because they are highly correlated; the latter particularly during afternoon hours. Second, to reduce the computational requirements, we use a stepwise algorithm to traverse the space of possible models. We start from the model $\mathcal{S} = \{1, \ldots, 6\}$ and $\mathcal{B} = \emptyset$, and in the forward selection step we successively include components in $\mathcal{B}$ that reduce *mAIC*. In the backward selection step we successively reduce the smooth components in set $\mathcal{S}$ (but only those which are not also a component in $\mathcal{B}$) if this reduces *mAIC*.

4 Empirical Analysis

4.1 Determinants of Bicycle Usage

We estimate the Poisson model for counts observed at all sites, and also for the four cumulative counts. Substantial over-dispersion was found throughout, particularly during

the evening peak period. For example, Table 2 reports the estimates of ϕ for *ALL*, which vary from a minimum of 3.4 at 06:00 to a maximum of 13.7 at 17:00. Table 2 also reports the estimated time trends for the four cumulative counts. There is a strong underlying upwards trend in usage throughout the period, particularly during the morning and evening commuting periods and for trips originating from the northern and inner city suburbs.

——Table 2 about here——

Table 3 summarises the models identified with the maximum *mAIC* values for the cumulative count *ALL* at each time of the day. For each univariate meteorological variable we denote the linear effects with 'L' and nonlinear effects with 'S'; non-zero bivariate interaction effects are denoted with 'B'. A substantial degree of parsimony is identified, with only a few of the bivariate interaction effects non-zero, and many of the univariate components linear. Similar parsimonious representations are also determined for the other count variables.

During the early morning peak (06:00 – 08:00) the effect of meteorology is well represented as additive in the six univariate components, with a nonlinear *MTEMP* effect throughout. Figure 4 plots the estimated meteorological effects for these three morning periods. The most important variables (as measured by function range) are *MTEMP*, *RAIN*, *RRAIN* and *WIND*, with the effects largely consistent during the three hours. Interestingly, it is not current air temperature, but daily maximum temperature, which usually occurs later between 14:00 and 17:00 in Melbourne, that proves important. In Section 5.2 we also show that there is strong dependence between cycling volumes during the morning and evening peak commuting periods. This is likely to be induced by strong dependence between mode of transport choice for the commuting trips to and from the workplace. Individuals therefore appear forward-looking in their decision-making and consider the forecast daily maximum temperature before selecting cycling as a mode of transport for the morning commute to work. A maximum temperature between 25 and 35$^{\mathrm{O}}$C is optimum, with less bicycle trips undertaken when *MTEMP* is outside this range. Higher levels of recent rainfall *RRAIN*

results in fewer cyclists; a similar effect is found with windspeed. Confidence bands are also given for the estimates at 06:00, which are similar to those for other times of the day. Estimates of the functional forms of *RAIN* and *RRAIN* have wide confidence intervals because of the small proportion of non-zero observations (9.9% and 17.8%, respectively).

During the evening peak period (16:00 – 19:00) the effects of current air temperature *TEMP* and recent rain *RRAIN* are both nonlinear and more pronounced than during the morning hours. This is likely because individuals are more backward-looking in their choice of mode of transport for the return commute. Figure 5 plots the nonlinear effects for the main evening commuting hour of 17:00, including the seasonal component. Recent rain has a substantial negative impact on propensity to cycle, whereas current rain has a linear coefficient with a p-value of 0.421; again, suggesting individuals are more backward-looking in their decision-making process. Windspeed is still important, but is now linear with a negative coefficient and p-value of 0.000.

Overall, while the impact of meteorology on bicycle commuting is complex, there are some main observations. Cyclists are deterred by maximum temperatures outside the comfortable region of 25–35^{o}C, high windspeeds and rainfall. However, in the morning they anticipate the weather conditions for the rest of the day because once they have made the decision to cycle to work, they are more likely to return by this mode of transport. In the evening, they are more backwards-looking and make the decision to cycle partially based upon the weather conditions revealed during the day.

——Table 3 and Figures 4 and 5 about here——

4.2 Cross-Price Elasticities

Table 4 provides the estimated elasticities $\hat{\alpha}_P$ for all loops at all times of the day. Four loops stand out as having significantly positive elasticities during the high volume commuting periods. These are loops 3, 7, 8 and 9, which are all inner city loops and include two of the

three highest volume loops in the data. This suggests that higher petrol prices are associated with an increase in bicycle trips recorded at these loops, presumably as commuters substitute from driving to cycling.

We note that loop 10 has negative elasticities during the commuting period. This loop is located on the beachfront 6.5km away from the CBD and is likely to have the highest proportion of recreational trips. These are likely to have quite different determinants of usage than the commuting trips that are predominantly counted by the other loops.

To examine further the spatial variation in elasticities, Table 5 provides the estimated elasticities for the cumulative counts at different times of the day. There is some evidence that commuters from the less affluent northern suburbs (N) are switching into cycling as a mode of transport as petrol prices rise, particularly in the early morning. However, there is stronger evidence that this is occuring during both the morning and evening commuting peaks with cyclists from wealthier suburbs (SE). The largest spatial variation in elasticities is associated with the distance of the loop from the city centre, with positive elasticities at all hours of the day for inner city (IC) loops that are significant throughout commuting hours. Presumably this is because cycling is a more attractive alterative to driving for trips to and from inner city suburbs and the CBD. This is consistent with Xing et al. (2010) who find shorter trip distance is a key factor for higher rates of adoption of bicycle transportation in six small US cities. On a city-wide basis there is a meaningful petrol price effect with all elasticities being positive for the cumulative count ALL, which are significant during the commuting periods. To benchmark these results, we also estimate a log-linear model with the same mean struture and in the same manner, but without Fisher scoring. The estimated elasticities are also found in Table 5 and they coincide with those from the Poisson model.

——Tables 4 and 5 about here——

4.3 Diagnostics

To assess the fitted models we employ the standardised Pearson residuals $\hat{\epsilon}_i = (y_i - \hat{\mu}_i)/\sqrt{\hat{\mu}_i\hat{\phi}}$ for $i = 1, \ldots, n$, where $\hat{\mu}_i$ and $\hat{\phi}$ are the point estimates from the best model $\mathcal{M}$ identified using *mAIC*. Following equation (3.3) the residuals should be approximately standard normal. Quantile plots suggest this to be the case and Table 6 contains p-values for the Kolmogorov-Smirnoff test of normality of the residuals for all cumulative counts at all times of day. The hypothesis that the residuals are standard normal can only be rejected at the 1% significance level in one of the 100 circumstances, suggesting the over-dispersed Poisson model is well-calibrated to the data.

To quantify the proportion of variation explained we use the statistic

$$R^2_{\text{full}} = 1 - \sum_{i=1}^{n} \hat{\epsilon}_i^2 \Big/ \sum_{i=1}^{n} \hat{\epsilon}_{0i}^2 \,,$$

where $\hat{\epsilon}_{0i}^2$ are the standardised Pearson residuals in the null model with intercept only. To assess the relative influence of different components we calculate the R^2 statistic using partial residuals on the numerator; that is, the residuals obtained from a fit when excluding specific components from η_i. The resulting full and partial R^2 statistics are shown in Figure 6 for the four cumulative counts where we exclude (i) trend and season; (ii) day type; (iii) petrol price; and (iv) meteorological components. The model explains around 80-90% of the variation in counts during peak hours, and 60-70% of variation during the midday period. The meteorological effects dominate, particularly during the midday period, followed by the seasonal and day type components. Petrol prices are the weakest of the four effects, and it would be difficult to determine its impact at an intraday level without controlling for the other effects.

——Table 6 and Figure 6 about here——

5 Intraday Serial Dependence

5.1 Copula Model

To account for intraday serial dependence in the counts we construct a multivariate model with $m = 14$ dimensions, each corresponding to an hourly count. We employ a Gaussian copula to capture dependence between the counts, and use the estimated hourly Poisson regression models as the marginal distributions. For hour j and observation i, an over-dispersed Poisson regression model with moments specified at equation (3.1) has a closed form probability mass function $\Pr(Y_{ij} = y_{ij}) = f_{ij}(y_{ij}; \mu_{ij}, \phi_j)$ given in Cameron & Trivedi (1986).[12] The joint distribution function of all 14 counts is

$$\begin{aligned} F(y_{i1}, \ldots, y_{im}) &= \Phi_m(\Phi_1^{-1}(u_{i1}), \ldots, \Phi_m^{-1}(u_{im}); C) \\ &= \Phi_m(y_{i1}^\star, \ldots, y_{im}^\star; C)\,. \end{aligned}$$

Here, Φ_m is the distribution function of an m-dimensional Gaussian distribution with zero mean and correlation matrix C, while Φ_1 is the distribution function of a standard univariate normal. We stress here that adopting a Gaussian copula model does not mean the counts are normally distributed; instead they follow a multivariate discrete distribution where the matrix C captures the dependence between the margins. As discussed in Danaher & Smith (2011), because the marginal distributions are discrete, $y_{ij}^\star$ is only known up to the bounds

$$\Phi_1^{-1}(F_{ij}(y_{ij} - 1; \mu_{ij}, \phi_j)) < y_{ij}^\star < \Phi_1^{-1}(F_{ij}(y_{ij}; \mu_{ij}, \phi_j))\,. \tag{5.1}$$

The bounds can be computed using the estimates of the marginal parameters μ_{ij}, ϕ_j, count data y_{ij} and the marginal distribution functions $F_{ij}(y_{ij}; \mu_{ij}, \phi_j) = \sum_{w=0}^{y_{ij}} f_{ij}(w; \mu_{ij}, \phi_j)$.

[12] Cameron & Trivedi (1986) define a specific Poisson-Gamma mixture model which they call 'negative binomial type 1' which has the moments of an over-dispersed Poisson model. This is not to be confused with the regular negative binomial model that Cameron & Trivedi call 'negative binomial type 2'.

Estimation of a Gaussian copula model with discrete margins is known to be a challenging problem (Song 2000). Direct maximum likelihood is computationally infeasible because evaluating the log-likelihood function is an $O(2^m)$ operation, which is impractical for $m = 14$. However, Pitt, Chan & Kohn (2006) and Danaher & Smith (2011) outline Bayesian analyses based on MCMC sampling schemes that treat $y^\star = \{y^\star_{ij}; i = 1, \ldots, n, j = 1, \ldots, m\}$ as latent and generate their values explicitly, along with the correlation matrix C. Such an approach is widely called 'Bayesian data augmentation', and is popular for the estimation of multivariate probit models (Chib & Greenberg 1998). We summarise the method in Appendix B, but refer the reader to Danaher & Smith (2011) for a complete discussion.

5.2 Empirical Analysis

We estimate the Gaussian copula for the cumulative count ALL in Table 1 using the Bayesian method discussed above. Because the copula model specifies the complete distribution F, a wide range of output can be obtained from the fitted parametric distribution to assess the level and impact of intraday serial dependence in the counts. Danaher & Smith (2011) outline how to compute pairwise Spearman's correlations $\rho^s_{j,k} = 12E(u_{ij}u_{ik}) - 3$ between counts y_{ij} and y_{ik}, where the expectation is evaluated with respect to the posterior distribution of C and $y^\star$. Figure 7(a) plots the Spearman correlations between all pairwise combinations of the hourly counts. Positive dependence $\rho^s_{j,j+1}$ can be seen between adjacent hours j and $j + 1$. This first order serial dependence is likely due to the omission from the model for μ_{ij} of additional determinants of usage that are themselves serially dependent. Interestingly, there is also positive dependence between trips in the morning and evening peaks periods. For example, the correlation between the peak hours of 08:00 ($j = 3$) and 17:00 ($j = 12$) is $\rho^s_{3,12} = 0.43$. This is evidence of a 'return trip' effect, where individuals who have already commuted to work by bicycle in the morning are more likely than otherwise would be the case to also return by bicycle, rather than switch to another mode of transport.

To illustrate the extent of this return trip effect, we compute the expected count during

the evening peak at 17:00, conditional upon that in the morning peak at 08:00. That is, $E(y_{12,i}|y_{3,i}) = \int E(y_{12,i}|y_{3,i}, y^\star, C)\pi(y^\star, C|\text{data})\text{d}(C, y^\star)$, where $(y^\star, C)$ is integrated out with respect to the posterior distribution $\pi(y^\star, C|\text{data})$. [13] Figure 7(b) plots this for 1 and 2 April 2008, and on both days there is a strong positive relationship between morning and expected evening counts.

The observed total daily counts for 1 and 2 April 2008 are plotted as vertical lines on Figure 7(d). Also plotted is the distribution of total daily counts $y_i^{\text{total}} = \sum_{j=1}^{14} y_{ij}$ from the fitted copula model for both days, and the actual observations are within the support of the distribution. On 1 April the total count was was high at 9649, but was much less on 2 April at a low of 4871. The reason is the advent of bad weather on 2 April, particularly with the strongest winds observed during the entire sample period. The average windspeed between 12:00 and 16:00 was 50.1km/h and was 38.4km/h during the entire 14 hours. Last, Figure 7(c) plots the posterior mean of the partial correlation matrix arising from C. These are the partial correlations of the latent variables $y_i^\star = (y_{i1}^\star, \ldots, y_{im}^\star)$ and again suggests strong first order intraday serial dependence.

6 Discussion

This paper presents an empirical study of bicycle commuting in a major western urban environment that is typical of many of today's car-dependent western cities. The data are unique in that they are both spatially disaggregate and, in particular, observed intraday. Previous studies are either based on survey data or highly aggregated with respect to time and/or location, which makes it hard to account for the impact of meteorology. This can result in the misnomer that weather conditions are a minor factor in the decision to cycle (Nankervis 1999; Hunt & Abraham 2007; Wardman et al. 2007), whereas our empirical work finds otherwise.

[13] This is evaluated in a Monte Carlo fashion by simulating 100 iterates from the copula model at each sweep of the sampling scheme, which is fast for a Gaussian copula; see Danaher & Smith (2011).

Our analysis confirms a strong diurnal variation in all aspects of the model, with sharp differences between the morning peak, evening peak and intermediate hours. We find that meteorological conditions are the strongest determinant of bicycle volumes throughout most of the day, followed by season and day type. The form of the effect is a complex multivariate function of six key meteorological variables, which are represented here using a parsimonious penalized spline decomposition. There is evidence that morning commuters are more forward-looking with regards to weather conditions, and evening commuters more backward-looking.

The extreme volatility in petrol prices between 2005 and 2008 also provides something akin to a natural experiment, allowing estimation of the cross-elasticities. In response to high petrol prices we find significant substitution into cycling as a mode of transport in the inner city, wealthy neighbourhoods and in the city overall, particularly during the peak commuting periods. As far as we are aware, the results in this study provide the first non-anecdotal evidence of such a phenomena in a major western urban environment. This relationship would be difficult to ascertain without controlling for season and weather conditions in a flexible fashion at an intraday resolution. The results suggest that policies aimed at encouraging bicycle commutes have a higher chance of success when the price of petrol is high, and in an inner city environment, where commuting trips are of shorter distance.

Using a Gaussian copula the model is extended to a multivariate model for longitudinal count data. Unsurprisingly, there is strong serial dependence in the hourly counts. More interestingly, even though the data is aggregate with respect to individuals, the results suggest a return trip effect at the individual level. If someone commutes to work in the morning by bicycle, then the late afternoon commute is more likely to also be by bicycle than would otherwise be the case. By fitting a copula model a range of additional inference is also available. This includes intraday forecasts of expected evening peak volumes, conditional upon those observed in the morning peak. Such forecasts have potential in transport management systems.

Last, we reiterate here that a two-stage estimator is used in Section 5. The first stage employs an approximate marginalized likelihood approach for estimating complex semiparametric count data models for each hour. The second stage is a Bayesian method to estimate a Gaussian copula to provide a 14-dimensional multivariate count data model. Because no strong priors are adopted, estimates based on the posterior distribution are very similar to those based only on the likelihood function, the latter of which cannot be employed directly because it is computationally intractable. Our approach is therefore analogous to two-stage maximum likelihood, which is a popular approach for the estimation of copula models (Cherubini et al. 2004). Fully Bayesian estimation, such as that employed in Herriges et al. (2008) of a linear demand system using multivariate count data, is difficult because of the complexity of also undertaking multivariate smoothing. To achieve this, it would be necessary to embed Bayesian multivariate smoothing methodology, such as that suggested by Wong & Kohn (1996), Smith & Kohn (1997), Lang & Brezger (2004) or Panagiotelis & Smith (2008), within the discrete-margined Gaussian copula framework, which is difficult computationally.

Acknowledgments

The work of Michael Smith was partially supported by Australian Research Council grant DP0985505, while Göran Kauermann acknowledges support provided by the Deutsche Forschungsgemeinschaft (DFG, project "Funktionale Modelle bei zeitstrukturierten Daten"). The authors would particularly like to thank VicRoads for providing the bicycle usage data. They would also like to thank Andrew John, Doug Dow and participants at a Melbourne Business School seminar for useful comments.

A Appendix

In this appendix we provide further details on how to compute the marginal likelihood and estimate the GLMM in Section 3.3. We write the log-likelihood contributions $l_i(\theta_{\mathcal{M}})$ in equation (3.9) in terms of the deviance (see McCullagh & Nelder, 1989, p. 197), so that $l_i(\theta_{\mathcal{M}}) = -d(y_i, \mu_i)/(2\phi)$, where $d(y, \mu) = -2\int_y^{\mu}(y-u)/u \, \mathrm{d}u$ and μ_i is the mean in equation (3.1). For a given model $\mathcal{M}$, we denote the concatenated design matrix of the nonlinear terms from the reduced rank bases as $B_{\mathcal{M}}$, and the corresponding concatenated basis coefficients as $v_{\mathcal{M}}$. The prior $v_{\mathcal{M}} \sim N(0, D_{\mathcal{M}}^{-1})$, where $D_{\mathcal{M}}$ is a block diagonal matrix with diagonal matrices $\lambda_s D_s$, $\lambda_j D_j, j \in \mathcal{S}$, and $\lambda_{jl} D_{jl}, (j,l) \in \mathcal{B}$ upon the leading diagonal. The logarithm of the marginal likelihood is

$$l_m(\alpha, \beta_{\mathcal{M}}, \lambda_{\mathcal{M}}; \mathcal{M}) = \log \int \exp\left\{ -\frac{1}{2\phi} \sum_i d(y_i, \mu_i) \right\} \psi(v_{\mathcal{M}}, D_{\mathcal{M}}^{-1}) \mathrm{d}v_{\mathcal{M}} \,, \tag{A.1}$$

where $\psi(x, V)$ denotes the multivariate normal density with zero mean, variance matrix V evaluated at point x. The integral in equation (A.1) is not tractable analytically, which has led to the use of the Laplace approximation; see Breslow & Clayton (1993). Rue et al. (2009) show that this Laplace approximation works well when computing posterior inference, while Kauermann et al. (2009) demonstrate the effectiveness of the Laplace approximation in penalized spline smoothing. Applying the Laplace approximation yields

$$l_m(\alpha, \beta_{\mathcal{M}}, \lambda_{\mathcal{M}}; \mathcal{M}) \approx -\frac{1}{2} \log |I + B_{\mathcal{M}}' M^{\star} B_{\mathcal{M}} D_{\mathcal{M}}^{-1}| - \frac{1}{2\phi} \sum_i d(y_i, \mu_i^{\star}) - \frac{1}{2} v_{\mathcal{M}}^{\star\prime} D_{\mathcal{M}} v_{\mathcal{M}}^{\star} \,. \tag{A.2}$$

Here, $M^{\star} = \mathrm{diag}(\mu_1^{\star}, \ldots, \mu_n^{\star})$, $\mu_i^{\star} = \exp(z_i'\alpha + \tilde{x}_i'\beta_{\mathcal{M}} + b_{\mathcal{M},i}' v_{\mathcal{M}}^{\star})$, $\tilde{x}_i$ are the linear terms from the basis expansions, $b_{\mathcal{M},i}'$ is the ith row of $B_{\mathcal{M}}$ and $v_{\mathcal{M}}^{\star}$ is the maximizer of the last two components in equation (A.2). The deviance is approximated using the squared Pearson residuals $(y_i - \mu_i^{\star})^2/\mu_i^{\star}$ as suggested in Pierce & Schafer (1986), see also McCullagh & Nelder (1989, p. 197). Note that $(y_i - \mu_i^{\star})/\mu_i^{\star} = \tilde{y}_i - \log(\mu_i^{\star})$, with $\tilde{y}_i$ the Fisher score defined in

equation (3.2). The Pearson residuals can be used to further approximate the (maximized) logarithm of the marginal likelihood (see Breslow & Clayton, 1993 or Wolfinger & O'Connell, 1993) as

$$l_m(\hat{\alpha}, \hat{\beta}_{\mathcal{M}}, \lambda_{\mathcal{M}}; \mathcal{M}) \approx -\frac{1}{2}\log|V_{\mathcal{M}}| - \frac{1}{2}(\tilde{y} - Z\hat{\alpha} - X\hat{\beta}_{\mathcal{M}})' V_{\mathcal{M}}^{-1}(\tilde{y} - Z\hat{\alpha} - X\hat{\beta}_{\mathcal{M}}), \qquad \text{(A.3)}$$

where $V_{\mathcal{M}} = (M^{\star -1} + B_{\mathcal{M}} D_{\mathcal{M}} B'_{\mathcal{M}})$, Z is a design matrix with ith row z_i and X is the design matrix for the linear terms in the basis decomposition of functions s and m. Maximization of equation (A.3) with respect to $\lambda_{\mathcal{M}}$ provides the final maximized value of the marginal likelihood.

B Appendix

In this section we outline the data augmentation approach used to estimate the Gaussian copula. This involves constructing a Markov chain Monte Carlo sampling scheme to evaluate the posterior distribution of the copula parameter matrix augmented with the latent variables, $\pi(C, y^{\star}|\text{data})$. We follow Danaher & Smith (2011) and employ the decomposition

$$C = \text{diag}(\Sigma)^{-1/2}\, \Sigma\, \text{diag}(\Sigma)^{-1/2},$$

where Σ is a positive definite matrix, and diag(Σ) a diagonal matrix comprised of the leading diagonal of Σ. We further decompose $\Sigma^{-1} = R'R$, with $R = \{r_{ij}\}$ being an upper triangular Cholesky factor with leading diagonal elements all ones. The sampling scheme is:

Sampling Scheme

Step 1. For $j = 1, ..., m$ and $i = 1, \ldots, n$:

Generate from $y^{\star}_{ij}|\{y^{\star} \backslash y^{\star}_{ij}\}, C, \text{data}$.

Step 2. For $i = 1, \ldots, m-1$, and $j > i$:

Generate from $r_{ij}|\{R \backslash r_{ij}\}, y^{\star}$ using random walk Metropolis-Hastings.

Step 3. Compute C from R.

The distribution in Step 1 is a constrained normal distribution, and is given in Danaher & Smith (2011), along with the Metropolis-Hastings acceptance ratio for Step 2. The sampling scheme is run for many sweeps, and after convergence, a Monte Carlo sample from the posterior distribution is obtained. From this sample parameter estimates and other posterior inference can be constructed.

References

Akaike, H., 1974. A new look at the statistical model identification. IEEE Transactions of Automatic Control, 19 (6), 716–723.

Australian Bureau of Statistics, 2005. National Regional Profile (by Statistical Local Area). viewed 1 Sept. 2009, `www.abs.gov.au`.

Austroads Inc., 2005. The Australian National Cycling Strategy 2005-2010. viewed 22 May 2009 at `www.vicroads.vic.gov.au`.

Bonham, J. & Suh, J., 2008. Pedalling the city: intra-urban difference in cycling for the journey-to-work. Road & Transport Research, 17 (4), 25–40.

Breslow, N.E. & Clayton, D.G., 1993. Approximate Inference in Generalized Linear Mixed Model. Journal of the American Statistical Association, 88 (421), 9–25.

Bhat, C.R. & Eluru, N., 2009. A copula-based approach to accomodate residential self-selection in travel behavior modeling. Transportation Research Part B, 43 (7), 749–765.

Cameron, A. & Trivedi, P., 1986. Econometrics Models Based on Count Data: Comparisons and Applications of Some Estimators and Tests. Journal of Applied Econometrics, 1 (1): 29–53.

Cherubini, U., Luciano, E., and Vecchiato, W., 2004. Copula Methods in Finance, Chichester, England: Wiley.

Cottet, R. & Smith, M., 2003. Bayesian Modeling and Forecasting of Intraday Electricity Load. Journal of the American Statistcal Association, 98 (464), 839–849.

Chib, S. & Greenberg, E., 1998. Analysis of multivariate probit models. Biometrika, 85 (2), 347–361.

Clarke, H. & Hawkins, A., 2006. Economic Framework for Melbourne Traffic Planning. Agenda, 13 (1), 63–80.

Danaher, P. & Smith, M., 2011. Modeling Multivariate Distributions using Copulas: Applications in Marketing (with discussion). Marketing Science, 30 (1), 4-21.

Eilers, P.H.C. & Marx, B.D., 1996. Flexible smoothing with B-splines and penalties. Statistical Science, 11 (2), 89–121.

Fiebig, D., Bartels, R. & Aigner, D., 1991. A Random Coefficient Approach to the Estimation of Residential End-Use Load Profiles. Journal of Econometrics, 50 (3), 297-327.

Ferdous, N., Eluru, N., Bhat, C.R. & Meloni, I., 2010. A multivariate ordered-response model system for adults' weekday activity episode generation by activity purpose and social context. Transportation Research Part B, 44 (8-9), 922–943.

Green, D.J. & Silverman, B.W., 1994. Nonparametric Regression and generalized linear models, London: Chapman & Hall.

Harville, D.A., 1977. Maximum likelihood approaches to variance component estimation and to related problems. Journal of the American Statistcal Association, 72 (358), 320–338.

Hastie, T., 1996. Pseudosplines. Journal of the Royal Statistical Society, Series B, 58 (2), 379-396.

Heinen, E., van Wee, B. & Maat, K., 2010. Commuting by Bicycle: An Overiew of the Literature. Transport Reviews, 30 (1), 59–96.

Herriges, J., Phaneuf, D. and Tobias, J., 2008. Estimating demand systems when outcomes are correlated counts. Journal of Econometrics, 147 (2), 282–298.

Hunt, J. & Abraham, J., 2007. Influences on bicycle use. Transportation, 34 (4), 453-470.

Kauermann, G., Krivobokova, T., & Fahrmeir, L., 2009. Some asymptotic results on generalized penalized spline smoothing. Journal of the Royal Statistical Society, Series B, 71 (2), 487–503.

Kauermann, G. & Opsomer, J.D., 2011. Data-driven Selection of the Spline Dimension in Penalized Spline Regression. Biometrika, 98 (1), 225–230. Lang, S. & Brezger, A., 2004. Bayesian P-Splines. Journal of Computational & Graphical Statistics, 13 (1), 183-212.

Lavergne, C., Martinez, M.-J. & Trottier, C., 2008. Empirical Model Selection in Generalized Linear Mixed Effects Models. Computational Statistics, 23 (1), 99–110.

McCullagh, P. & Nelder, J.A., 1989. Generalized Linear Models, 2nd edition, London: Chapman and Hall.

Nankervis, M., 1999. The effect of weather and climate on bicycle commuting. Transportation Research Part A, 33 (6), 417–431.

O'Sullivan, F., 1988. Nonparametric estimation of relative risk using splines and cross-validation. SIAM Journal on Scientific Computing, 9 (3), 531–542.

Panagiotelis, A. & Smith, M., 2008. Bayesian identification, selection and estimation of semiparametric functions in high-dimensional additive models. Journal of Econometrics, 143 (2), 291–316.

Pierce, D.A. & Schafer, D.W., 1986. Residuals in generalized linear models. Journal of the American Statististical Association, 81 (396), 977–986.

Pitt, M., Chan, D. & Kohn, R., 2006. Efficient Bayesian Inference for Gaussian Copula Regression Models. Biometrika, 93 (3), 537–554.

Pucher, J., Dill, J. & Handy, S., 2010. Infrastructure, programs and policies to increase bicycling: An international review. Preventive Medicine, 50, Supplement 1, S106–S125.

Rietveld, P. & Daniel, V., 2004. Determinants of bicycle use: do municipal policies matter?. Transportation Research Part A, 38 (7), 531–550.

Rue, H. Martino, S. & Chopin, N., 2009. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. Journal of the Royal Statistical Society, Series B, 71 (2), 1–35.

Ruppert, D., 2002. Selecting the number of knots for penalized splines. Journal of Computational & Graphical Statistics, 11 (4), 735–757.

Ruppert, D., Wand, M. & Carroll, R., 2009. Semiparametric regression during 2003-2007. Electronic Journal of Statistics, 3, 1193-1256.

Ruppert, D., Wand, M. & R. Carroll, R., 2003. Semiparametric Regression, Cambridge: Cambridge University Press.

Sener, I. N., Eluru, B. & Bhat, C.R., 2010. On jointly analyzing the physical activity participation levels of individuals in a family unit using a multivariate copula framework. Journal of Choice Modelling, 3 (3), 1–38.

Shannon, T., Giles-Corti, B., Pikora, T., Bulsara, M., Shilton, T. & Bull, F., 2006. Active commuting in a university setting: Assessing commuting habits and potential for modal change. Transport Policy, 13 (3), 240–253.

Shively, T. & Sager, T., 1999. A semiparametric regression approach to adjusting for meteorological variables in air pollution trends. Environmental Science & Technology, 33 (21), 3873–3880.

Silva, R., and Lopes, H., 2008. Copulas, Marginal Distributions and Model Selection: A Bayesian Note. Statistics and Computing, 18 (3), 313–320.

Smith, M. & Kohn, R., 1997. A Bayesian approach to nonparametric bivariate regression. Journal of the American Statistical Association, 92 (440), 1522–1535.

Song, P. X.-K., 2000. Multivariate Dispersion Model Generated from Gaussian Copula. Scandinavian Journal of Statistics, 27 (2), 305–320.

Stinson, M. & Bhat, C.R., 2004. Frequency of bicycle commuting: internet-based survey analysis. Transportation Research Record, 1878, 122–130.

Vaida, F. & Blanchard, S., 2005. Conditional Akaike information for mixed effects models. Biometrika, 92 (2), 351–370.

VicRoads, 2001. Cycling to Work in Melbourne 1976-2001. viewed on 22 May 2009 at `www.vicroads.vic.gov.au`.

Wager, C., Vaida, F. & Kauermann, G., 2007. Model selection for p-spline smoothing using Akaike information criteria. Australian & New Zealand Journal of Statistics, 49 (2), 173–190.

Wahba, G., 1978. Improper priors, spline smoothing and the problem of guarding against model errors in regression. Journal of the Royal Statistical Society, Series B, 40 (3), 364–372.

Wahba, G. 1990. Spline Models for Observational Data. CBMS-NSF Regional Conference Series in Applied Mathematics, Vol. 59, Phildelphia, SIAM.

Wardman, M., Tight, M. & Page, M., 2007. Factors influencing the propensity to cycle to work. Transportation Research Part A, 41 (4), 339–350.

Wand, M. 2003. Smoothing and mixed models. Computational Statistics, 18, 223–249.

Wecker, W. & Ansley, C., 1983. The Signal Extraction Approach to Nonlinear Regression and Spline Smoothing. Journal of the American Statistical Association, 78 (381), 81–89.

Wedderburn, R.W.M., 1974. Quasi-likelihood functions, generalized linear models, and the Gauss-Newton method. Biometrika, 61 (3), 439–447.

Wolfinger, R., & O'Connell, M., 1993. Generalized linear mixed models: A pseudo-likelihood approach. Journal of Statistical Computation & Simulation, 48 (3), 233–243.

Wong, C. & Kohn, R., 1996. A Bayesian approach to additive semiparametric regression. Journal of Econometrics, 74 (2), 209–235.

Wood, S., 2003. Thin plate regression splines. Journal of the Royal Statistical Society, Series B, 65 (1), 95–114.

Wood, S. 2006. Generalized Additive Models, London: Chapman & Hall.

Xing, Y., Handy, S. and Mokhtarian, P., 2010. Factors associated with proportions and miles of bicycling for transportation and recreation in six small US cities. Transportation Research Part D, 15 (2), 73–81.

	Loop Characteristics			Average Hourly Counts			Time Frame and Available Data			
Loop No	Induction Loop Name	Distance to GPO	Suburb Location	06:00−09:00	16:00−19:00	06:00−19:00	Number of Workdays	Start Date	End Date	Percent Missing
1	St. Georges Road (Sumner Road)	5.3km	North	87.87	87.52	60.79	550	1 Dec 05	17 Jun 08	11.6
2	Anniversary Trail 1	9.0km	East	15.59	15.08	14.69	584	12 Dec 05	17 Jun 08	5.3
3	Main Yarra Trail (North Bank)	2.8km	East	209.15	201.46	141.77	570	13 Dec 05	19 Jun 08	7.6
4	Main Yarra Trail (South Bank)	2.9km	South	93.84	90.62	64.61	564	12 Dec 05	19 Jun 08	8.7
5	Canning St. (Carlton)	2.6km	North	164.46	157.14	111.10	587	1 Dec 05	19 Jun 08	6.1
6	Upfield Railway Line	4.1km	North	59.28	49.75	40.24	585	12 Dec 05	19 Jun 08	5.3
7	Capital City Trail (Princes Hill)	3.8km	North	75.57	63.79	50.57	578	1 Dec 05	19 Jun 08	7.5
8	Footscray Road Path	1.7km	West	135.94	146.43	93.92	549	12 Dec 05	20 Jun 08	11.1
9	Tram 109 Trail	2.3km	South	64.40	58.98	45.08	581	12 Dec 05	18 Jun 08	5.8
10	Bay Trail (St. Kilda)	6.5km	South	69.29	73.62	60.18	587	1 Dec 05	18 Jun 08	6.0
	Cumulative Counts over Multiple Loops									
N	North (Loops 5-7)			300.0	271.4	202.3	576	21 Dec 05	19 Jun 08	5.8
SE	South & East (Loops 2-3,9,10)			358.6	348.7	261.5	570	13 Dec 05	17 Jun 08	9.1
IC	Inner City (Loops 3,5-7,9)			574.3	532.6	389.7	556	23 Dec 05	18 Jun 08	8.7
ALL	All Locations (Loops 2,3,5-7,9,10)			659.0	620.5	464.2	554	23 Dec 05	17 Jun 08	9.1

Table 1: Location and spatial characteristics of inductive loop counters, average working day hourly counts, and the time frame and data availability for working days. Summaries are given for all ten loops, and also for cumulative counts over multiple loops. The percentage of missing is relative to the length of data, and an observation is considered missing for a cumulative count when one is missing at any component loop.

	Hour													
	6	7	8	9	10	11	12	13	14	15	16	17	18	19
	Estimated Trend Coefficients for Cumulative Counts													
SE	0.092**	0.086**	0.081**	0.031**	0.008	-0.02	-0.05*	-0.08**	-0.13**	-0.11**	-0.02	0.093**	0.019	-0.02
N	0.163**	0.177**	0.158**	0.111**	0.088**	0.094**	0.052**	0.055**	0.056**	0.138**	0.111**	0.129**	0.124**	0.128**
IC	0.157**	0.139**	0.125**	0.083**	0.076**	0.048*	0.018	-0.02	-0.05	0.028	0.044*	0.112**	0.083**	0.082**
ALL	0.118**	0.117**	0.120**	0.069**	0.044**	0.019	-0.01	-0.03*	-0.06**	0.003	0.031*	0.105**	0.061**	-0.69*
	Estimated Over-Dispersion Parameter ϕ for ALL													
	3.50	8.20	9.45	4.53	7.56	8.91	8.74	8.61	11.7	12.0	12.4	13.6	11.3	6.24

Table 2: Estimated linear trends (annualized) for the logarithm of the mean count $\eta_i = \log(E(y_i))$ defined in Section 3.1. Values significant at the 5% level are denoted with '$*$', while those significant after a Bonferroni adjustment for the 14 hours of the day are denoted with '$**$'. Also reported are the estimates of the over-dispersion parameter ϕ for the cumulative count *ALL*.

	Hour													
	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Univariate Effects														
TEMP	L	L	L	L	S	S	L	L	L	L	S	S	S	S
MTEMP	S	S	S	S	S	S	S	S	S	S	L	S	S	S
RAIN	S	S	S	S	S	L	S	L	L	S	L	L	L	L
RRAIN	L	L	L	L	L	S	L	L	L	L	S	S	S	S
HMD	L	L	L	L	L	S	S	S	L	L	L	S	L	L
WIND	L	S	S	S	S	S	S	S	S	S	L	L	L	S
Bivariate Interaction Effects														
TEMP & WIND	.	.	.	.	.	.	.	.	.	.	.	.	.	B
MTEMP & RAIN	.	.	.	.	B	.	.	.	.	.	.	.	.	.
MTEMP & HMD	.	.	.	.	.	B	.	.	.	.	.	.	.	.
MTEMP & WIND	.	.	.	B	B	B	.	.	B	B	.	.	.	.
RAIN & WIND	.	.	.	.	.	.	B	.	.	.	.	.	.	.
HMD & WIND	.	.	.	.	.	.	B	B	.	.	.	.	.	.

Table 3: Meteorological effects selected for the cumulative count *ALL*. Univariate effects labelled with 'L' or 'S' were identified as linear or nonlinear, respectively. Bivariate interaction effects identified as zero are labelled with '.', and non-zero with 'B'. Omitted bivariate interactions were zero at all times of the day.

	Hour													
	Morning Peak (06:00-09:00)												*Evening Peak (16:00-19:00)*	
Loop No	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	0.42*	0.07	0.16	-0.41*	0.09	0.95**	0.33	0.11	0.56*	0.81**	0.81**	0.3*	0.13	0.09
2	0.95*	0.21	-1.17**	-0.29	0.11	0.15	0.26	-0.98*	-0.49	-1.3**	-0.19	0.02	0.53	-0.53
3	0.65**	0.85**	0.82**	0.65**	-0.1	0.59	0.26	0.5	0.97*	1.22*	1.12**	0.9**	0.89**	0.82**
4	-0.47*	-0.02	0.32**	-0.27	0.04	-0.53	-0.24	-0.17	-0.15	-0.27	-0.02	0.03	0.26*	0.39*
5	0.32*	0.07	-0.02	0.09	0.1	0.34*	0.38*	-0.04	0.46*	0.18	0.38**	0.09	0.01	0.08
6	0.46*	0.05	-0.17	0.06	-0.21	0.26	0.3	0.4	-0.15	-0.47*	-0.48*	0.12	0.12	-0.11
7	1.25**	0.39**	0.24*	0.48**	0.12	0.18	0.1	0.23	0.47	0.13	0.21	0.04	-0.42*	-0.15
8	0.51**	0.56**	0.16	0.34*	0.1	0.35	-0.02	0.44	0.32	0.93**	0.4*	0.44**	0.82**	-0.05
9	-0.22	0.56**	0.21*	0.39*	0.52	0.41	0.58*	0.06	0.51*	0.1	0.87**	0.56**	0.44**	0.47
10	-0.8**	-1.24**	-0.68**	-0.15	0.13	0.34	0.22	-0.27	0.22	0.23	-0.39	-0.73**	-0.41*	-0.76**

Table 4: Estimates of α_P, the cross-elasticity with respect to petrol prices, using the Poisson model for all ten loops and times of the day. Values significant at the 5% level are denoted with '$*$', while those significant after a Bonferroni adjustment for the 14 hours of the day are denoted with '$**$'.

	Hour													
	Morning Peak (06:00-09:00)										*Evening Peak (16:00-19:00)*			
	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Count	*Poisson Model*													
N	0.65**	0.16	0.02	0.19*	0.03	0.25	0.26	0.14	0.28	-0.01	0.15	0.09	-0.09	0.01
SE	0.11	0.39**	0.35**	0.31*	0.04	0.37	0.39	0.06	0.46	0.39	0.66**	0.5**	0.54**	0.19
IC	0.48**	0.55**	0.31**	0.29**	0.08	0.38*	0.38*	0.29	0.59*	0.43*	0.62**	0.49**	0.39**	0.34**
ALL	0.31*	0.33**	0.21*	0.24*	0.04	0.34*	0.28	0.06	0.4*	0.22	0.45**	0.38**	0.3*	0.08
Count	*Log-Linear Model*													
N	0.69**	0.23*	0.07	0.15	0.04	0.22	0.34*	0.22	0.32*	0.01	0.19	0.13	-0.06	0.02
SE	0.13	0.35**	0.34**	0.4**	0.27	0.41	0.48*	0.09	0.47	0.29	0.61**	0.54**	0.49**	0.18
IC	0.48**	0.56**	0.31**	0.28**	0.16	0.33*	0.45*	0.39*	0.56**	0.39*	0.59**	0.49**	0.36**	0.32*
ALL	0.26*	0.33**	0.18*	0.25*	0.1	0.36*	0.38*	0.16	0.41*	0.23	0.48**	0.34**	0.26*	0.14

Table 5: Estimates of the cross-elasticity with respect to petrol prices using the Poisson model for the four cumulative counts. Values significant at the 5% level are denoted with '$*$', while those significant after a Bonferroni adjustment for the 14 hours of the day are denoted with '$**$'.

	Hour													
Location	6	7	8	9	10	11	12	13	14	15	16	17	18	19
N	0.995	0.789	0.593	0.874	0.693	0.874	0.090	0.989	0.693	0.834	0.939	0.120	0.078	0.742
SE	0.629	0.866	0.976	0.824	0.130	0.530	0.195	0.130	0.150	0.006	0.019	0.222	0.530	0.778
IC	0.932	0.626	0.391	0.864	0.148	0.677	0.018	0.434	0.083	0.022	0.052	0.071	0.083	0.901
ALL	0.900	0.573	0.524	0.820	0.624	0.900	0.110	0.674	0.167	0.217	0.146	0.146	0.127	0.862

Table 6: The p-values for Kolmogorov-Smirnoff test of normality for the standarized Pearson residuals.

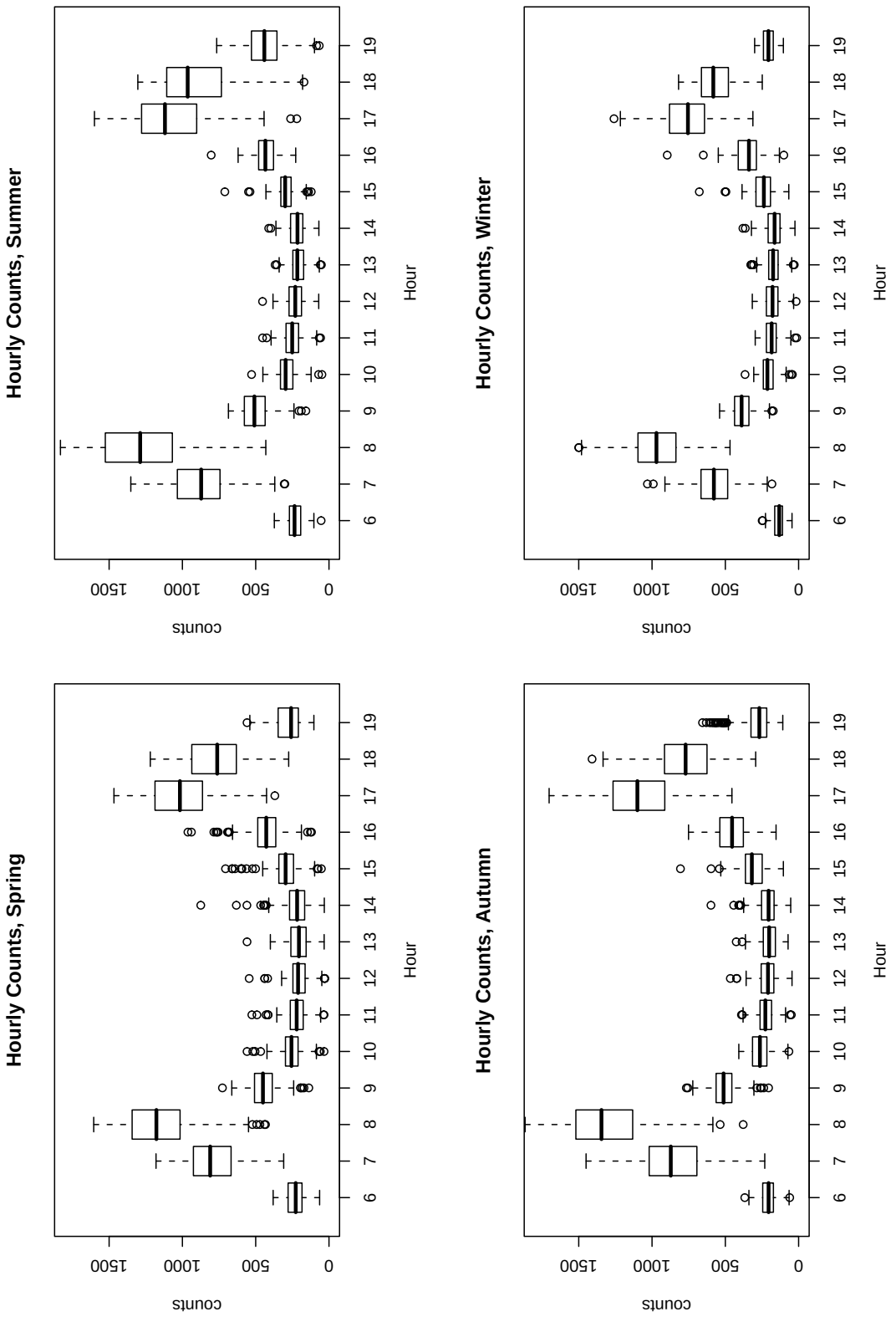

Figure 1: Hourly city-wide count (*ALL*) broken down by both time of day and southern hemisphere season.

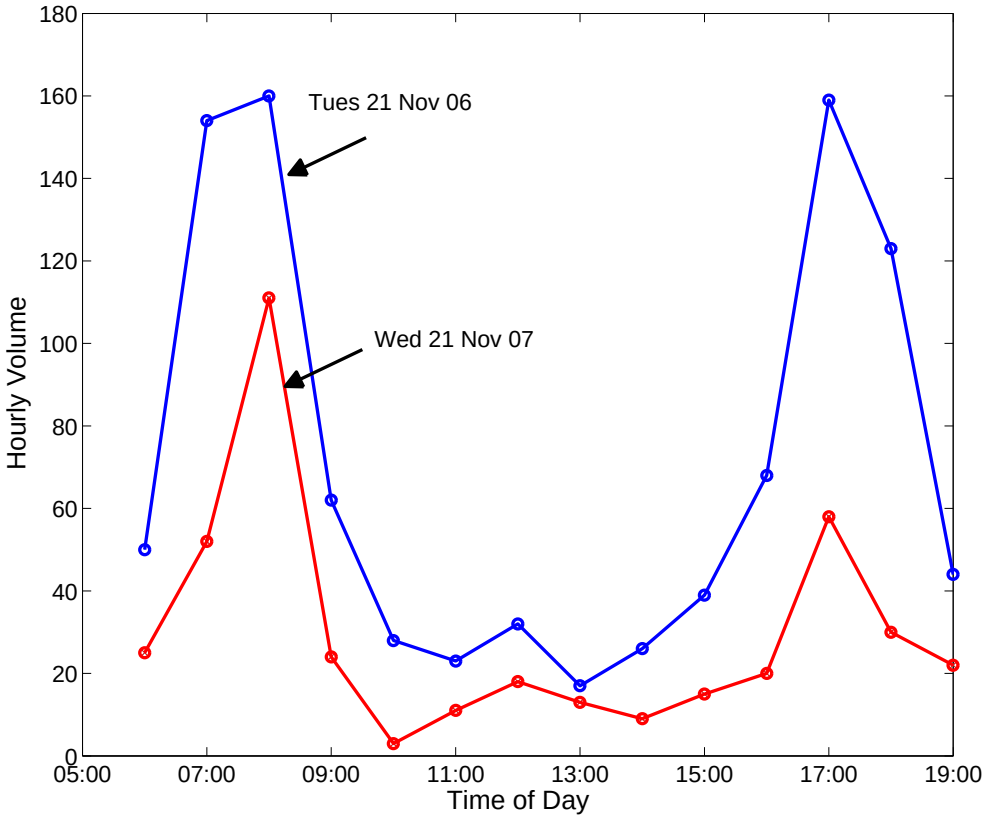

Figure 2: Plot of hourly volumes at the St Georges Road inductive loop between 06:00 and 19:00. The blue line depicts the counts on 21 November 2006, while the red line depicts the counts on 21 November 2007.

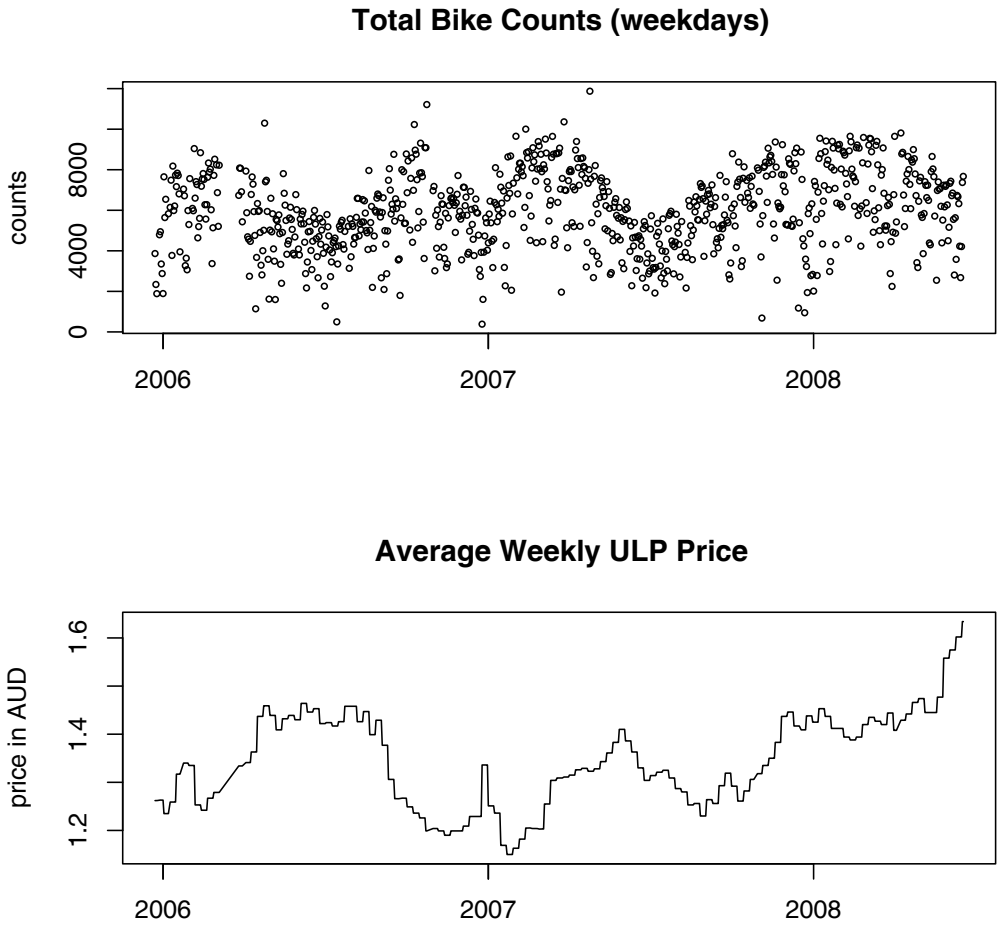

Figure 3: Total daily counts across all loops and average metro petrol prices in Melbourne between December 2005 and June 2008.

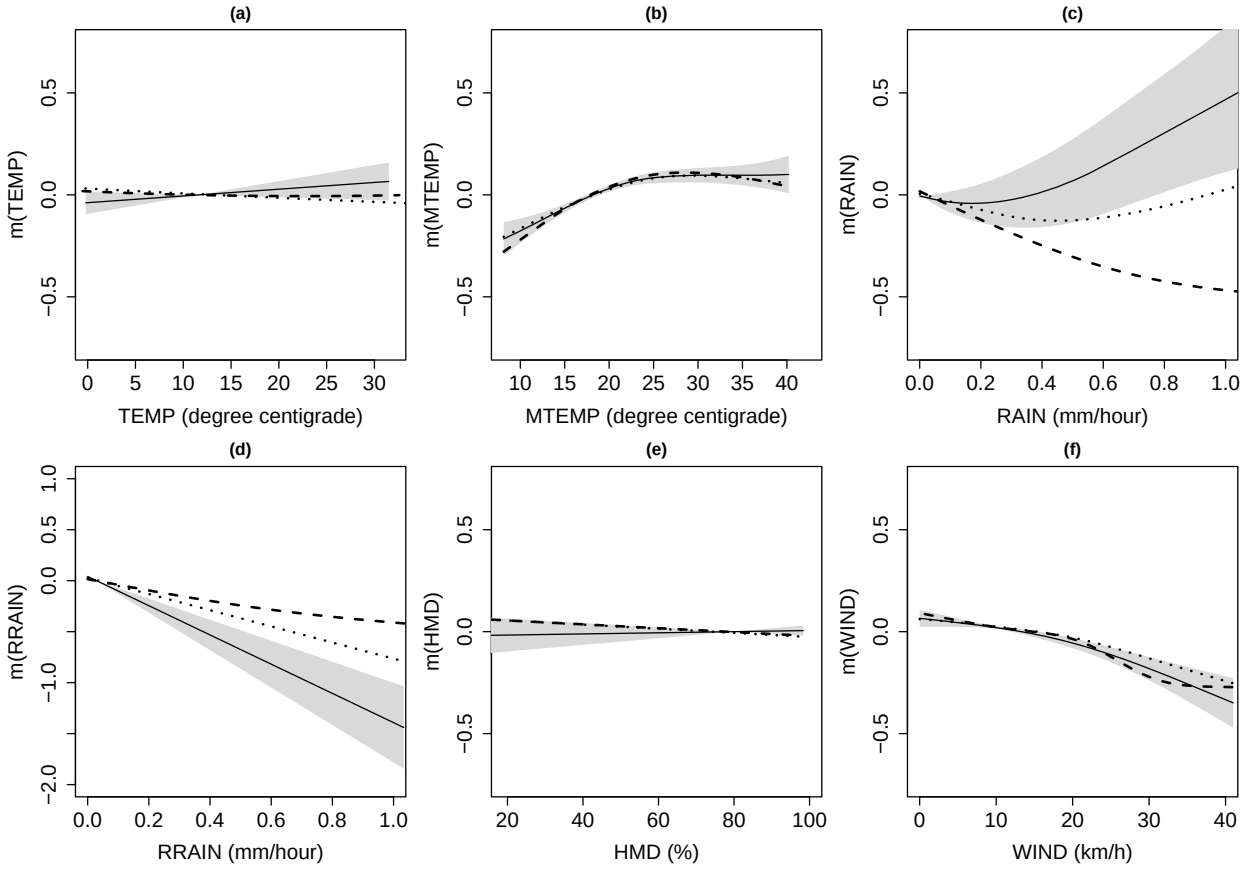

Figure 4: Estimated meteorological effects for the cumulative count *ALL* during the morning peak period. The estimates are for 06:00 (solid line), 07:00 (dashed lines) and 08:00 (dotted lines). The shaded areas are 95% confidence bands for the estimates at 06:00.

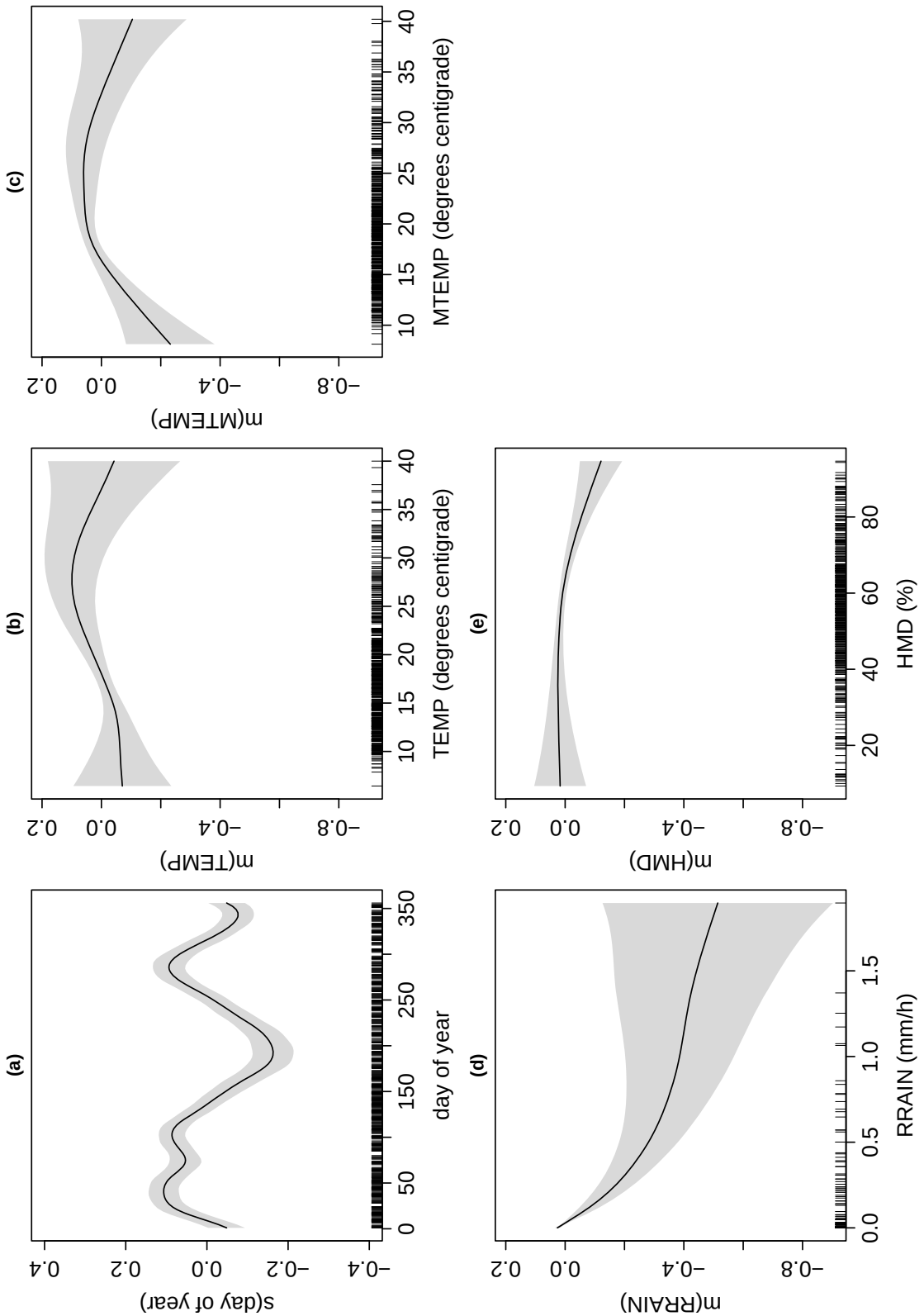

Figure 5: Estimated nonlinear effects for the cumulative count *ALL* at 17:00, along with 95% confidence bands for the function estimates. The distribution of the independent variables are indicated on the horizontal axes.

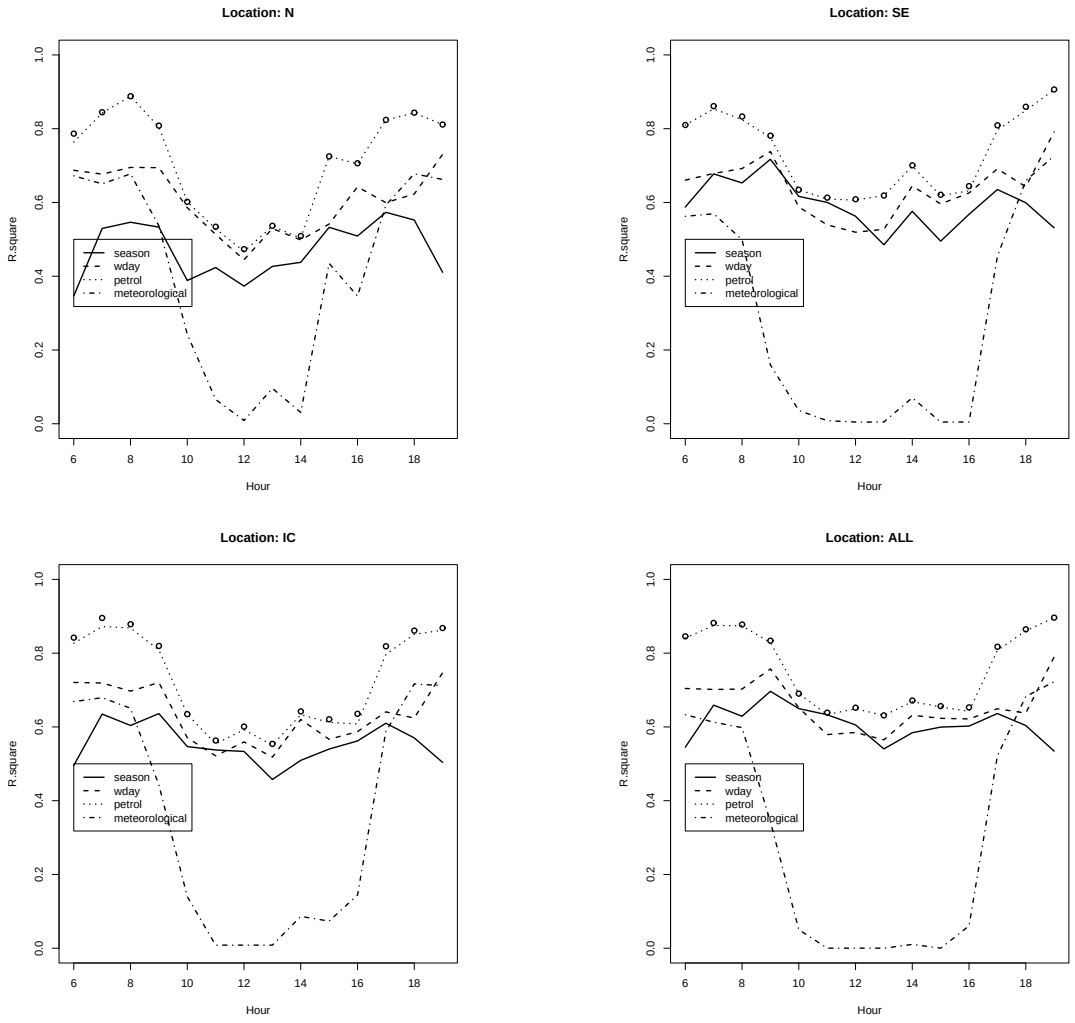

Figure 6: R^2 plots for the four cumulative counts *N*, *SE*, *IC* and *ALL* against time of day. The statistic R^2_{full} is plotted as hollow circles, while the partial R^2 values are given as different line types for models excluding (i) seasonal component s (bold); (ii) day of week dummies (long dashed); (iii) meteorological effects m (dot-dashed); and, petrol price term α_P (small dots).

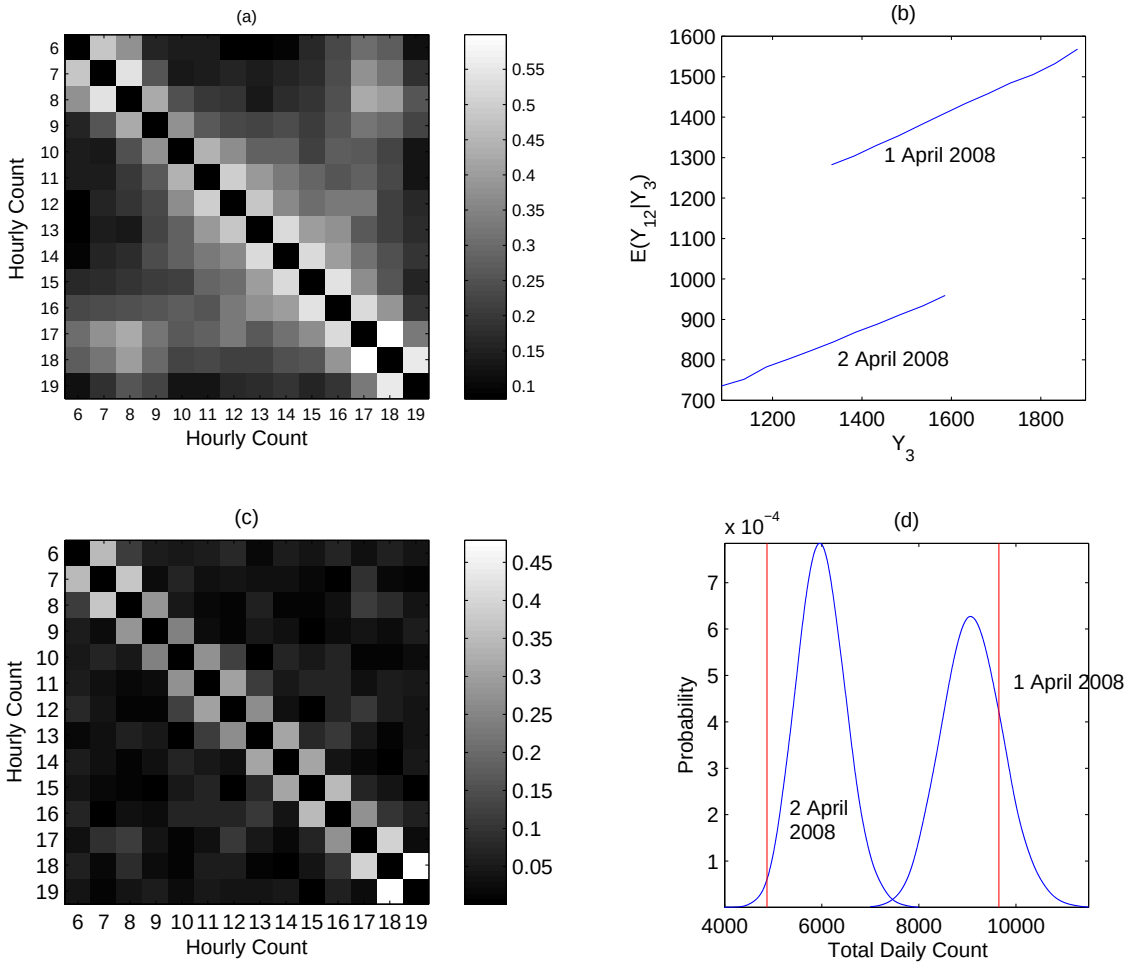

Figure 7: Bayesian posterior inference from the fitted Gaussian copula model, with Poisson regression margins, for the 14 hourly observations of the cumulative count *ALL*. Panel (a): matrix of pairwise Spearman correlations $\rho^s_{j,k}$. Panel (b): expected evening peak count at 17:00 ($j = 12$), conditional upon the morning peak count at 08:00 ($j = 3$). Panel (c): partial correlations from the copula parameter matrix C. Panel (d): distribution from the copula model of total daily counts, y_i^{total}, on April 1 and 2, 2008.