Obesity and labour market outcomes in Italy: a dynamic panel data evidence with correlated random effects

Antonio Pacifico

Received: 23 November 2020 / Accepted: 21 June 2022 / Published online: 22 July 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract
This paper investigates the effects of obesity, socio-economic variables, and individual-specific factors on work productivity across Italian regions. A dynamic panel data with correlated random effects is used to jointly deal with incidental parameters, endogeneity issues, and functional forms of misspecification. Methodologically, a hierarchical semiparametric Bayesian approach is involved in shrinking high dimensional model classes, and then obtaining a subset of potential predictors affecting outcomes. Monte Carlo designs are addressed to construct exact posterior distributions and then perform accurate forecasts. Cross-sectional Heterogeneity is modelled nonparametrically allowing for correlation between heterogeneous parameters and initial conditions as well as individual-specific regressors. Prevention policies and strategies to handle health and labour market prospects are also discussed.

Keywords Bayesian inference · MCMC algorithms · Density forecasts · Heterogeneous effects · Healthcare Statistics · Work productivity

JEL classification A1 · C01 · E02 · H3 · N01 · O4

Introduction
This paper aims to improve the existing literature on the relationships between obesity, healthcare, and labour market outcomes using a dynamic panel data (DPD) with correlated random effects (CRE). The main thrust of this study is to examine in depth these relationships across Italian regions using: (i) additional time-varying socio-economic indicators and exogenous factors to deal with endogeneity issues (because of unobserved heterogeneity and omitted variables) and functional forms of misspecification; (ii) unobserved time-invariant individual-specific heterogeneity affecting outcomes; (iii) semiparametric Bayesian inference to avoid the curse of dimensionality and variable selection problems in high dimensional settings; and (iv) conditional density forecasts performing policy-relevant strategies.

Obesity is a complex condition increasing the risk of various health problems, such as cancer, diabetes, hypertension, and depression, which may negatively affect economic outcomes and thus the probability of employment (see, e.g., [13, 23, 58, 67], and [4]). Furthermore, employers may discriminate against obese workers, resulting in worse labor market performance (see, e.g., [11, 36], and [51]). In that context, it is not always clear as discrimination may be directly related to labour market outcomes such as social participation, per-hour wage rates, sales (see, e.g., [17] and [9]). Earlier studies highlight that physical appearance may be directly related to productivity, affecting the individual’s capacity to work. For instance, obese people may be related to non-desirable personality traits potentially affecting productivity, becoming less socially and intellectually skilled than their non-obese counterparts (see, e.g., [60, 61, 66], and [69]).

The previous literature has primarily looked at the effects of obesity on wages that would be not necessarily the same as the ones of obesity on employment. However, if wages are not fully flexible or obesity causes serious health problems, there would be a severe causal-effect of obesity on employment (see, for instance, [1]). In this context, investigating the effect of obesity on labour market outcomes is complicated due to potential reversed causality and endogeneity issues. Moreover, there may also be unobserved heterogeneous
characteristics that vary systematically between obese and non-obese people, affecting employment. For instance, whether high discount rates can encourage people to be more prone to weight-gaining consumption, they would make investments in human capital and thus future labour market outcomes less attractive ([20, 21] and [22]).

Important issues of the effects of obesity on labour market outcomes can be further addressed focusing on groups having traditionally lower participation rates such as females. The existing literature on the effects between obesity and wage productivity is still limited, reaching mixed conclusions. In [20, 21] and [55], no effect of obesity on employment among US men and women is obtained. On the other hand, in [53, 54], a significant and negative effect for both genders in the UK are found. Different results are found in Italy instead. Even if the country is characterized by one of the slimmest adult populations in Europe (OECD 2017), Italy is experiencing a strong increase in the portion of obese people, mainly among children age, making it a particularly relevant problem to be addressed (see, e.g., [16]). Moreover, because Italy has always been the country of fashion, larger attention is devoted to physical appearance. Thus, according to international studies, weight-based discrimination would be a relevant topic, strongly affecting every aspect of obese people’s everyday life including labour market outcomes (see, e.g., [18] and [65]). However, since obesity denotes a major issue in every developed country, the existing literature has devoted much less attention to the weight bias in the job market compared to race and gender discrimination (see, for instance, [2]).

The relationship between high body weight and labour market outcomes has been primarily studied using data from developed and high income countries, such as the US and West Europe (e.g., England, Denmark, and Finland). In the last decade, the main studies concerning labour market outcomes have focused on wages/earnings, employment, and occupational selection (see, e.g., [28, 14, 15, 44, 62, 29, 42], and [30]). Related papers, focused on the US, have used the National Longitudinal Survey of Youth (NLSY) data, founding mixed results (see, e.g., [63, 50], and [57]). Some shortcomings of these studies are that they ignore potential endogeneity and spatial interdependence of obesity accounting for small and unrepresentative samples and cross-sectional data (see, for instance, [26] and [34]).

Concerning DPD models, they are widely used in empirical economics for forecasting individuals’ future outcomes (see, e.g., [33, 38, 46], and [48]), and allowing the possibility of controlling for unobserved time-invariant individual heterogeneity (see, e.g., [11] and [7] for linear case). Such heterogeneity is an important issue and failure to control for it results in misleading inferences. That problem is even more severe when the unobserved heterogeneity may be correlated with covariates. However, in these traditional methods, with large cross-sectional dimension (N) and short (fixed) time-series (T), the estimators of homogenous parameters would result biased and inconsistent due to incidental parameter problems (because of individual-specific effects). Indeed, leaving the individual heterogeneity unrestricted, the number of individual-specific effects would grow with the sample size. Moreover, with short T, the estimates of the heterogeneous parameters would be highly contaminated from idiosyncratic shocks obtaining inaccurate forecasts.

The methodology proposed in this paper focuses on the aforementioned issues and consists of a hierarchical semiparametric Bayesian approach for inference in dynamic panel setup with cross-sectional heterogeneity. Methodologically, Conjugate Informative Proper Mixture (CIPM) priors are used to obtain a subset of best combination of predictors (or model solutions) affecting outcomes of interest, where best stands for the model solution providing the most accurate predictive performance over all candidate models. Markov Chain Monte Carlo (MCMC) algorithms and implementations are used to construct exact posterior distributions, and then perform better density forecasts and policy issues.

The contributions of this paper are threefold. First, full Bayesian methods are used to deal with potential features in real-world data such as density forecasts, incidental parameters, endogeneity issues, and structural model uncertainty (because of one or more parameters are posited as the source of model misspecification problems). CIPM priors and MCMC-based Posterior Model Probabilities (PMPs) are used to: (i) include cross-sectional individual information from the whole panel, acting as a strong model selection in high dimensional model classes; (ii) outline assumptions and specification strategy of the informative priors on covariates, outcomes of interest, and the distribution of unobserved heterogeneity; and (iii) deal with overfitting and model uncertainty when addressing variable selection problems. Here, ‘overfitting’ refers to the overestimation of effect sizes since more complex models (or models with a larger number of unknown parameters) always provide a somewhat better fit to the data than simpler models.

Second, I build on and extend the [56]’s analysis who develops a Robust Open Bayesian (ROB) procedure in two stages for implementing Bayesian Model Averaging (BMA) and Bayesian Model Selection (BMS) in multiple high dimensional linear regression models. More precisely, I extend the prior specification strategy in a panel context with either dynamic or time-invariant factors, taking the name of Time-varying ROB (TROB) procedure. The main novelty of the implementation is the use of nonparametric Bayesian inference for heterogeneous
parameters. Conversely, homogeneous parameters are modelled through parametric priors. In the first stage, a subset of covariates better explaining and thus fitting the data are chosen according to their PMPs. The latter denotes the probability of each candidate model performing the data. In the second stage, further shrinkage is performed to obtain a smallest final subset of best submodels having statistically significant predictive capability. Finally, the submodel with the highest Bayes Factor in logarithm (IBF) would correspond to the final solution containing a potential subset of candidate predictors with a higher significant overall $\overline{R}^2$ (\(\overline{\hat{R}}^2\)) measure, where ‘strong’ refers to $\overline{R}^2$ value equal to or bigger than 30%.

Third, MCMC algorithms and implementations are addressed to construct exact posterior distributions dealing with semiparametric forecasting problems. Better density forecasts are involved in DPB with CREs because of three main features: (i) the use of a hierarchical full Bayesian approach; (ii) the framework of informative mixture priors; and (iii) the observation of incidental parameters.

An empirical example across Italian regions describes the estimating procedure and forecasting performance. It covers the period 2006 – 2018, where conditional density forecasts are performed extending it up to 2021 to also address potential findings concerning the impact of the ongoing pandemic disease on the global economy. Then, the sample is divided into three groups to better highlight the strengths and limits of policy strategies: (i) Northern Italy; (2) Central Italy; and (3) Southern Italy. In this study, natural conjugate priors are used to deal with persistent structural breaks affecting the panel data analysis: 2007, due to the global financial crisis; 2011, due to fiscal consolidations in Northern Italy; (2) Central Italy; and (3) Southern Italy. In this study, natural conjugate priors are used to deal with persistent structural breaks affecting the panel data analysis: 2007, due to the global financial crisis; 2011, due to fiscal consolidations in post-financial crisis; and 2014, due to recovery fund.

The remainder of this paper is organized as follows. Section "Econometric method and estimation procedure" introduces the econometric model, the Bayesian inference, and the estimating procedure. Section "Multivariate dynamic panel data" describes the dynamic procedure to outline prior specification strategy, and then construct posterior distributions. Section "Bayesian analysis" displays theoretical properties concerning dynamic panel data and the semiparametric Bayesian approach. Section "Dynamic analysis" describes the data, the empirical analysis, and density forecasts along with policy issues. The final section contains some concluding remarks.

**Econometric method and estimation procedure**

**Multivariate dynamic panel data**

The baseline dynamic panel data (DPD) model with correlated random effects (CREs) and cross-sectional heterogeneity is:

\[ y_{it} = \lambda_i + \beta y_{i,t-1} + \alpha x_{it} + \gamma z_{i,t-1} + \epsilon_{it}, \]  

(1)

where \( i = 1, 2, ..., N, \ t = 1, 2, ..., T, \ y_{it} \) is a \( N \times 1 \) vector of outcomes, \( y_{i,t-1} \) and \( z_{i,t-1} \) are \( N \times 1 \) vectors of predetermined and endogenous variables for each \( i \), respectively, with \( \beta \) and \( \gamma \) denoting autoregressive coefficients to be estimated for each \( i \), \( \lambda_i \) is a \( N \times 1 \) vector of strictly exogenous factors for each \( i \), with \( \alpha \) denoting the regression coefficients to be estimated, \( \alpha_i \) is a \( N \times 1 \) heterogeneous intercept containing, for example, time-constant differences among regions, and \( \epsilon_{it} \sim i.i.d. N(0, \sigma_{\epsilon}^2) \) is a \( N \times 1 \) vector of unpredictable shock (or idiosyncratic error term), with \( E(u_{it}) = 0 \) and \( E(u_{i1} \cdot u_{is}) = \sigma_{u}^2 \) if \( i = j \) and \( t = s \), and \( E(u_{i1} \cdot u_{is}) = 0 \) otherwise.

In this study, for notational simplicity, I assume only one lag for every time-varying parameters. Here, some considerations are in order: (i) the predetermined variables refer to the lagged outcomes \((y_{i,t-1})\) containing, for example, control variables; (ii) the \( \lambda_i \) 's denote cross-sectional heterogeneity affecting the outcomes \( y_{it} \); (iii) CREs matter and then the \( \lambda_i \) 's are treated as random variables and possibly correlated with some of the covariates within the panel; (iv) the strictly exogenous factors \( x_{it} \) contain dummy variables to test, for example, the presence of policy shifts (or structural breaks); and (v) the instruments are fitted values from AutoRegressive (AR) parameters based on time-varying lagged variables.

Furthermore, because of the location aspects in Eq. (1), two potential problems estimating the regression coefficients in Eq. (1) need to be accounted for: spatial dependence among observations in the system and spatial heterogeneity among time-varying parameters in the relationships. Both of them lead to violations of assumptions underlying unbiasedness, consistency, and efficiency (see, for instance, [10] and [40]). According to Eq. (1), spatial interdependence is involved in the estimation including AR coefficients, and spatial heterogeneity is addressed letting the model be multidimensional (panel data).

Assuming stationarity to hold in Eq. (1), the time-series regressions are valid and the estimates feasible. Thus, some moment restrictions need to hold to address exact identification in a context of CREs, and then estimate the time-varying parameters (\( \beta \) and \( \gamma \)) for \( T \geq 3 \) (see, for instance, [3, 8], and [12]). More precisely, I assume that \( \lambda_i \) and \( \epsilon_{it} \) are independently distributed across \( i \) and have the familiar error components structure:

\[ E(\lambda_i) = 0, E(\epsilon_{it}) = 0, E(\epsilon_{it} \cdot \lambda_i) = 0 \]  

\[ \text{for} \quad i = 1, \ldots, N \quad \text{and} \quad t = 2, \ldots, T, \]  

(2)

and

\[ E(\epsilon_{it} \cdot \epsilon_{is}) = 0 \]  

\[ \text{for} \quad i = 1, \ldots, N \quad \text{and} \quad t \neq s. \]  

(3)

Then, I also assume the standard assumption concerning the initial conditions \( y_{it=1} \).
\[ E(y_{i,t+1} \cdot u_{it}) = 0 \quad \text{for} \quad i = 1, \ldots, N \quad \text{and} \quad t = 2, \ldots, T. \]  

(4)

**Bayesian analysis**

The main thrust of the TROB procedure is threefold. First, it performs a strong model selection to obtain the best model solution better explaining and thus fitting the data in high dimensional panel setups. Second, CIPM prior allow for modelling nonparametrically common (\( \beta, \alpha, \gamma \)) and heterogeneous (\( \lambda_i, \sigma^2 \)) coefficients, and then their distribution changing in a corresponding fashion in accordance with different model solutions. Third, MCMC algorithms and implementations can be constructed to perform better density forecasts concerning solutions. Third, MCMC algorithms and implementations can be constructed to perform better density forecasts concerning solutions. Then, I combine the homogeneous parameters into a strong model selection to obtain the best model and the parameter space at the same time. The shrinking process jointly deals with overfitting (or overestimation of effect sizes) and model uncertainty (involved in the procedure) by posterior model probabilities (PMPs) for every candidate model. They are defined as:

\[ \pi \left( y | \theta_j \right) = \int_B \pi \left( y, \lambda_i | \theta_j, M_j \right) \cdot d\lambda, \]  

(8)

where \( B \) denotes the multidimensional (natural) parameter space for \( \theta_j, M_j = (M_1, \ldots, M_m) \) denotes a countable collection of all (potential) model solutions for any individual \( i \). The integrand in Eq. (8) is defined as:

\[ \int_B \pi \left( y, \lambda_i | \theta_j, M_j \right) = \pi \left( \theta_j, \lambda_i, M_j | y \right) \cdot \pi \left( y | M_j \right), \]  

(9)

The error terms (\( e_{it} \)) are individual-time-specific shocks characterized by zero mean and homoskedastic Gaussian innovations. In a hierarchical framework, I combine the individual-specific heterogeneity into the vector \( \varphi_i = \left( \lambda_i, \sigma^2 \right) \) under cross-sectional heterogeneity and homoskedasticity. Assuming correlated random coefficients model, \( \varphi_i \) and \( s_{it} \) may be correlated with each other, with:

\[ s_{it} = \left( y_{i,t}^\varphi, y_{i,t}^\varepsilon, x_{i0:t}, z_{i,t-1}, \cdot \cdot \cdot \right), \]  

(5)

and

\[ D \left( y_{it} | y_{i,t-1}, x_{it}, z_{i,t-1}, \lambda_i \right) = D \left( y_{it} | y_{i,t-1}, x_{it}, z_{i,t-1}, y_{it} \cdot \lambda_i \right). \]  

(6)

The error terms (\( e_{it} \)) is individual-time-specific shocks characterized by zero mean and homoskedastic Gaussian innovations. In a hierarchical framework, I combine the individual-specific heterogeneity into the vector \( \varphi_i = \left( \lambda_i, \sigma^2 \right) \) under cross-sectional heterogeneity and homoskedasticity. Assuming correlated random coefficients model, \( \varphi_i \) and \( s_{it} \) may be correlated with each other, with:

\[ s_{it} = \left( y_{i,t}^\varphi, y_{i,t}^\varepsilon, x_{i0:t}, z_{i,t-1}, \cdot \cdot \cdot \right). \]  

(7)

Given these specifications, the DPD model in Eq. (1) tends to be less parsimonious and suitable for density forecasting. Thus, I perform a fully Bayesian approach to estimate the prior distribution of the correlated random effects by pooling the cross-sectional information from the whole panel. The main difference between an empirical and fully Bayesian approach is that the former picks the \( \lambda_i \)’s distribution by maximizing the maximum likelihood of the data (see, e.g., [25, 38, 45, 41], and [32, 33]). Conversely, fully Bayesian methods construct a prior for the correlated random effects and then evaluates it in view of the observed panel data (see, for instance, [46] and [48] (linear case); and [47] (non-linear case)). In addition, the fully Bayesian approach tends to be more suitable for density forecasting and more easily for extending the TROB procedure to the nonlinear case.

Let \( \mathcal{F} \) be the full panel set containing all (potential) model solutions, the variable selection problem is addressed by imposing an auxiliary indicator variable \( \theta_j \), with \( j = 1, 2, \ldots, m \), containing every possible \( 2^m \) subset choices, where \( \theta_j = 0 \) if \( \theta_j \) is small (absence of \( j \)-th covariate in the model) and \( \theta_j = 1 \) if \( \theta_j \) is sufficiently large (presence of \( j \)-th covariate in the model). According to the [56]’s framework, I run the TROB procedure by pooling all high dimensional model solutions (\( m \geq 15 \)) to shrink the model and the parameter space at the same time. The shrinking process jointly deals with overfitting (or overestimation of effect sizes) and model uncertainty (involved in the procedure) by posterior model probabilities (PMPs) for every candidate models.
where $M_k$ denotes the submodel solutions of the DPD in Eq. (1), with $M_k < M_j$, $k \leq j$, $\{1 \leq k < j\}$, and $\tau$ is a threshold chosen arbitrarily for an enough posterior consistency\(^2\).

Because of the panel considered in this study features high dimensional cross-sectional units $N$ (predictors $\geq 15$) and non-fixed time-series $T$, I use $\tau = 0.5\%$ to jointly manage all equations within the system (through the conditioning set $s_{it}$), and their potential interactions (through time-varying common covariates).

The second stage entails reducing the model space $S$ and then obtaining a smaller subset of best submodel solutions:

$$S = \left\{ M_k : M_k \subset S, S \in \mathcal{F}, \Delta_k \subset \Delta, \sum_{k=1}^{}\pi \left(M_k | y_i = y_i, \theta \right) \geq \tau \right\},$$

(10)

where $M_k \subset M_k$, $\pi (M_k | y_i = y_i, \theta)$ denotes the PMPs, with $\theta$ referring to a new auxiliary variable containing the only best model solutions in the subset $S$ and $\tau$ referring to a new arbitrary threshold to evaluate the probability of the model solutions in $S$ performing the data (PMPs). In this study, I still use $0.5\%$ for $\tau$, independently of $N$, for a sufficient prediction accuracy in explaining the data.

The final model solution ($M_{e*}$) consists of a subset of the submodel solutions obtained in the second stage ($M_k \subset E$) to perform conditional density forecasting and policy-making. It will correspond to the outcome with higher log natural Bayes Factor (IBF):

$$lBF_{e*} = \log \left\{ \frac{\pi (M_{e*} | y_i = y_i)}{\pi (M_{k} | y_i = y_i)} \right\}.$$

(12)

In this study, the IBF is interpreted according to the scale evidence in [56], but with more stringent conditions:

$$\begin{align*}
0.00 < lBF_{e*} & \leq 4.99 \quad \text{no evidence for submodel } M_{e*} \\
5.00 < lBF_{e*} & \leq 9.99 \quad \text{moderate evidence for submodel } M_{e*} \\
10.00 < lBF_{e*} & \leq 14.99 \quad \text{strong evidence for submodel } M_{e*} \\
lBF_{e*} > 15.00 & \quad \text{very strong evidence for submodel } M_{e*}.
\end{align*}$$

(13)

\(^2\) In Bayesian analysis, posterior consistency ensures that the posterior probability (or PMP) concentrates on the true model.

### Dynamic analysis

#### Prior specification strategy and Bayesian predictor

The variable selection procedure entails estimating $\theta$, and $\theta$ as posterior means (the probability that a variable is in the model). All observed variables in $s_{it}$ and individual heterogeneity $\varphi_i$ are hierarchically modelled via CIPM priors:

$$\pi (\theta, \varphi, \theta) = \pi (\theta | \theta) \cdot \pi (\lambda_i | \theta, s_{0i}) \cdot \pi (\sigma_i^2 | \theta, s_{0i}) \cdot \pi (\theta),$$

(14)

where

$$\pi \left( \theta | \mathcal{Y}_{-1} \right) = N \left( \bar{\theta}_{i|t}, \bar{\theta}_{i|t} \right),$$

(15)

$$\pi (s_{0i} | \lambda_i) = N (0, \xi),$$

(16)

$$\pi (\theta) = w_{|\theta|} \cdot \left( \frac{j}{|\theta|} \right)^{-1},$$

(17)

$$\pi (\sigma_i^2) = IG \left( \frac{\omega_0}{2}, \frac{v_0}{2} \right),$$

(18)

with $N(\cdot)$ and $IG(\cdot)$ standing for Normal and independent Inverse-Gamma distributions, respectively, $\mathcal{Y}_{-1}$ referring to the cross-sectional information available at time $t - 1$, and $w_{|\theta|}$ denoting the model prior choice related to the sum of the PMPs (or Prior Inclusion Probabilities) with respect to the model size $|\theta|$, through which the $\theta$’s will require a non-0 estimate or the $\theta$’s should be included in the model. In this way, one would weigh more according to model size and, setting $w_{|\theta|}$ large for smaller $|\theta|$, assign more weight to parsimonious models. Moreover, the use of independent $IG(\cdot)$ distribution allows cross-equation independence of the coefficient distributions and remove the dependence of $s_{it}$ and $\varphi_i$ (see, for instance, [27]).

All hyperparameters are known. More precisely, collecting them in a vector $\phi$, where $\phi = \left( \bar{\theta}_{i|t}, \bar{\theta}_{i|t}, \xi, w_{|\theta|}, \omega_0, v_0 \right)$, they are treated as fixed and are either obtained from the data to tune the prior to the specific applications (such as $\bar{\theta}_{i|t}, w_{|\theta|}, \omega_0$) or selected a priori to produce relatively loose priors (such as $\bar{\theta}_{i|t}, \xi, v_0$). Here, $w_{|\theta|}$ is restricted to a benchmark prior $\max \left( nT, |\theta| \right)$ according to the non-0 components of $\theta$.

Nevertheless, to accommodate the correlated random coefficients model where the individual-specific heterogeneity ($\lambda_i$) can be correlated with the conditioning variables $s_{0i}$ and then $y_{0i}$, I use a non-parametric Bayesian prior accommodating conditional density forecasting. As suggested in [59], I consider the Mixture of Gaussian Linear Regressions (MGLR)
prior and extend it to the multidimensional setup conditioning on \( s_{i0} \):

\[
\lambda_i | s_{i0} \sim \sum_{b=1}^{\infty} p_b(s_{i0}) N \left( \mu_b \left[ 1, s_{i0}^T \right] , \Omega_b \right),
\]

(19)

where \( b = 1, 2, \ldots \) is a component label, \((\mu_b, \Omega_b) \sim G_b\) are component parameters, with \( G_b \) denoting a conjugate multivariate-normal-inverse-gamma (or multivariate-normal-inverse-Wishart) distribution for scalar \( \lambda_i \), and \( p_b \) is a component probability. The mixture probabilities are characterized by a Probit Stick-Breaking Process:

\[
p_b(s_{i0}) = \Phi \left( \zeta_b(s_{i0}) \right) \Pi_{j \neq b} \left( 1 - \Phi \left( \zeta_j(s_{i0}) \right) \right),
\]

(20)

where \( \zeta_b \) is a common prior, with \( \zeta_b \sim N(0, 1) \) as in [64]. This prior allows to borrow information across populations by shrinking the stick-breaking ratios that corresponds to the \( l \)-th mixture component towards a common value, with \( l \) denoting the distribution of \( \psi_i \). However, this formulation would hold for random effect models, and then I implement it by evaluating \( \zeta_b \) as a stochastic function drawn from Gaussian process with zero mean and variance \( V_b \), where \( V_b = \sigma^2 \cdot I_k \) in homoskedastic case. In this way, I am able to estimate it by drawing posterior distributions for \( \sigma^2 \).

With these specifications, four considerations are in order: (i) component means are linear in \( s_{i0} \); (ii) component covariances are independent of \( s_{i0} \); (iii) mixture probabilities can be modelled as functions of \( s_{i0} \); and (iv) CIPM priors involved in TROB procedure absorb dependency on \( s_{i0} \), without requiring any further dependency of component means and covariances on \( s_{i0} \) beyond the MGLR specification. This latter consists of a methodological contribution with respect to the [59]‘s framework according to high dimensional panel settings.

In the compound decision theory, the infeasible optimal oracle (or benchmark) predictor implies that \( \psi_i \) and the distribution of the unobserved heterogeneity \((\pi(\lambda_i, s_{i0}))\) are known and the values of \( \lambda_i \) are unknown across units. Thus, I construct a feasible semiparametric Bayesian predictor to provide an estimate of this distribution and then perform better conditional density forecasts asymptotically converging to the oracle forecast. The density forecasts of \( y_{i,T+m} \) are defined as:

\[
l_{i,T+m}^{\text{cond}}(y|\theta, l, F_i) = \int \pi(\psi_i|\theta, l, F_i) \cdot \pi(\psi_i|\theta, l, F_i) d\psi_i,
\]

(21)

where \( l_{i,T+m}^{\text{cond}}(\cdot) \) denotes the predictor conditional on the common \((\theta)\) and heterogeneous \((\psi_i)\) parameters, with \( F_i \) referring to the individual \( i \)'s data. In Eq. (21), the first term captures individual \( i \)'s uncertainty due to future shock \( s_{iT} \), with \( s_{iT} = (y_{i,0:T}, x_{i,0:T}, \sigma^2_{i,0:T}, \sigma_{i,0:T}, \sigma_{i,0:T}) \), and the second term accounts for individual-specific posteriors corresponding to:

\[
\pi(\psi_i|\theta, l, F_i) = \frac{\prod_{t=1}^{T} \pi(y_i|\psi_i, \theta, s_{i,t-1}) \cdot I_l(\psi_i) d\psi_i}{\int \prod_{t=1}^{T} \pi(y_i|\psi_i, \theta, s_{i,t-1}) I_l(\psi_i) d\psi_i},
\]

(22)

with \( s_{i,t-1} = (y_{i,0:t-1}, x_{i,0:t-1}, \sigma^2_{i,0:T}, \sigma_{i,0:T}, \sigma_{i,0:T}) \). By construction, dealing with high dimensional time-series, I am able to extend the density forecasts in Eq. (21) to \( \sigma \)-period-ahead forecasts by pooling the information from the whole panel. In this way, I am able to update all the prior beliefs to construct posterior distributions of \((\theta, l)\), and then the feasible semiparametric Bayesian predictor for \( \sigma \)-period-ahead forecasts \((l_{i,T+m}^{\text{cond}}(\cdot))\) by integrating the conditional predictor over the posteriors. This latter is defined as:

\[
l_{i,T+m}^{\text{cond}}(y|\theta, l, F_i) = \int l_{i,T+m}^{\text{cond}}(y|\theta, l, F_i) \cdot d\pi(\theta, l|F_i) d\theta dl,
\]

(23)

where, according to the main goal of this study, the first term reflects functional forms of misspecification due to future shocks on multiple-period-ahead forecasts, and the second term deals with endogeneity issues due to cross-sectional individual-specific heterogeneity in high dimensional dynamic panel setups. Methodologically, the infeasible oracle predictor (or benchmark) would correspond to the semiparametric Bayesian predictor in Eq. (23) by plugging the true (expected) values of the common parameters \((\theta_0)\) and underlying distribution of \( \psi_i (l_0) \) into the conditional predictor in Eq. (21):

\[
l_{i,T+m}^{\text{oracle}}(y|\theta, l, F_i) = l_{i,T+m}^{\text{cond}}(y|\theta_0, l_0, F_i).
\]

(24)

**MCMC algorithms and density forecasts**

Despite the dramatic parameter reduction involved in the TROB procedure, the analytical computation of posterior distributions of \((\theta, l|y_{i,T+m})\) is unfeasible, where \( y_{i,T+m} \) denotes the expectations of outcomes associated with density forecasts [Eq. (21)]. Thus, I use MCMC implementations to draw conditional posterior distributions of \((\theta_1, \theta_2, \ldots, \theta_T|y_{i,T+m})\). More precisely, I include a variant of the Gibbs sampler approach, the Kalman-Filter technique, to analytically obtain forward recursions for posterior means and covariances on \( s_{iT} \). They are, respectively:

\[
\tilde{\theta}_i[l] = \tilde{\theta} + Z_i[l-1] + \left[ (y_i - (s_i[l] F_i) \tilde{\theta}_i[l-1]) \cdot P_i[l-1]^{-1} \right],
\]

(25)

\[
\tilde{\theta}_i[l] = \left[ I_k - \left( P_i[l-1] \cdot (s_i[l] F_i)^T \right) \right] \cdot \tilde{\theta}_{i-1}[l-1].
\]

(26)
with
\[ P_{t|t-1} = \left( \alpha_t | F_t \right)^{\gamma} \cdot P_{t|t-1} \cdot \varepsilon_t + \sigma^2. \]  
(27)

Thus, the marginal distributions of \( \theta \) can be computed by averaging over draws in the nuisance dimensions, and the Kalman filter backward can be run to obtain their posterior distributions:
\[ \pi(\theta_t|\theta_{t-1}, \gamma_{t,T+m}, \tilde{\theta}_{t-1}) = N\left( \tilde{\theta}_{t|T+m}^{\gamma}, \hat{\theta}_{t|T+m} \right), \]  
(28)

with
\[ \hat{\theta}_{t|T+m} = \left( \tilde{\theta}_{t|T+m}^{\gamma} \cdot \hat{\theta}_{t|t} \right) + \left( \sum_{t=1}^{T} (s_{it}^{\gamma} \cdot (\sigma^{\gamma})^{-1} \cdot (s_{it}^{\gamma} \cdot F_{it}) \right), \]  
(29)

\[ \tilde{\theta}_{t|T+m} = \left[ I_k - \left( \tilde{\theta}_{t|T+m}^{\gamma} \cdot \hat{\theta}_{t|t} \right) \right] \cdot (\tilde{\theta}_{t|t}), \]  
(30)

where \( \tilde{\theta}_{t|T+m}^{\gamma} \) and \( \hat{\theta}_{t|T+m} \) are smooth \( \sigma \)-period-ahead forecasts of \( \theta \) and covariances of the forecast error. The above output of the Kalman filter is used to generate a random trajectory for \( \{ \theta_t \}_{t=1}^{T} \) by the backward recursions on \( s_{it}^{\gamma} \), starting with a draw of \( \{ \theta_t \} \) from \( N\left( \tilde{\theta}_{T|T}^{\gamma}, \hat{\theta}_{T|T} \right) \).

Given Eq. (28), the other posterior distributions can be defined as:
\[ \pi(\delta_{it}) = N(0, \tilde{\xi}), \]  
(31)

\[ \pi(\theta) = \tilde{w}_{|\theta|} \left( \frac{\kappa}{|\theta|} \right)^{-\frac{1}{2}}, \]  
(32)

\[ \pi(\delta^2) = IG\left( \frac{\tilde{\omega}}{2}, \tilde{v} \right), \]  
(33)

Here, some considerations are in order. In Eq. (31), \( \tilde{\xi} = \xi + V_{i,i}(l) \), with \( \xi \) and \( V_{i,i}(l) \) denoting the arbitrary scale parameter and the posterior predictive variance of \( \lambda \) according to its distribution \( l \) in homoskedastic case, respectively. In this analysis, \( \xi \approx 1.0 \) and \( V_{i,i}(l) = \delta^2 \cdot I_k \) according to the sample size \( |\theta| \) as described in Eq. (15).

In Eq. (32), \( \tilde{w}_{|\theta|} \) refers to the model posterior choice according to the sum of the PMPs obtained in the second stage with respect to model size \(|\theta|\), with \( \tilde{w}_{|\theta|} = max(N_{TT}, |\theta|) \) accounting for the non-0 components in \( \mathcal{E} \).

In Eq. (33), \( \tilde{\omega} = \omega_0 \cdot \kappa \) and \( \tilde{v} = v_0 \cdot \kappa \), with \( \omega_0 \) and \( v_0 \) denoting the arbitrary degrees of freedom (sufficiently small) and the arbitrary scale parameter, respectively. In this analysis, \( \omega_0 \approx 0.1 \) and \( v_0 \approx 0.001 \).

Finally, the last two common parameters to be defined are \( \mu_{\theta} \sim N(0, \delta) \) and \( \Omega_{\theta} \sim IW(c_1, \psi_1) \). In conjugate settings and flexible mixture probabilities involved in CIPM priors, I use another MCMC integration, the Gibbs sampling, to accommodate the underlying distribution \( l \) conditional on \( s_{it}^{\gamma} \). Methodologically, this framework would be similar to approximating the conditional density of individual-specific heterogeneity through Bayes’ theorem, but without explicitly modelling the distribution of \( l(\varphi_i|s_{it}) \). In this way, the resulting feasible predictor in (23) would be still close to the oracle one, and then both the estimated common parameters \( (\hat{\beta}, \hat{a}, \hat{\gamma}) \) and the estimated distribution of the cross-sectional heterogeneous effects \( \hat{\lambda}(\lambda_i|s_{it}) \) achieve posterior consistency. According to the MGLR prior in Eq. (19), the posteriors are:
\[ \pi(\mu_{\theta}) = N(0, \delta), \]  
(34)

\[ \pi(\Omega_{\theta}) = IW(c_1, \psi_1). \]  
(35)

where \( \hat{c}_1 = c_1 \cdot k, \hat{\delta} = \delta \cdot k, \) and \( \psi_1 = \psi_1 + (c_1 \cdot c_2) \), with \( c_1 \) denoting the arbitrary degrees of freedom (sufficiently small), and \( \delta \) and \( \psi_1 \) referring to the arbitrary scale parameters. In this analysis, \( c_1 \approx 0.1, \delta \approx 0.001, \) and \( \psi_1 \approx 0.001 \).

The accuracy of the density forecasts is measured by the Predictive Score in logarithm (lPS) as in [31]:
\[ lPS = \frac{1}{N} \sum_{i} \log\left( \hat{\pi}(y_{i,t+1}|F_i) \right), \]  
(36)

where \( \hat{\pi}(y_{i,t+1}|F_i) \) denotes the predictive likelihood with respect to the estimated model conditional on the observed data \( F_i \).

**Theoretical properties**

The proof of the posterior consistency is as follows.

**Assumption 4.1** (Identification: General Model) Consider the DPD model in Eq. (1):

1. **Model Setup**
   a. \( (s_{it}, \lambda_i, \sigma^2) \) are i.i.d. across \( i \).
   b. For all \( t \), conditional on \( (y_{it}, s_{it-1}, \gamma_{it}) \), \( y_{it}^c \) is independent of \( \varphi_i = (\lambda_i, \sigma^2) \).
   c. \( x_{i,t;T} \) is independent of \( \varphi_i = (\lambda_i, \sigma^2) \).
   d. Conditioning on \( (s_{it}, \lambda_i) \) and \( \sigma^2 \), they are independent of each other.
   e. Let \( e_i \sim N(0, \sigma^2) \) is i.i.d. across \( i \) and independent of \( (s_{it-1}, \lambda_i) \).

---

3 See, for instance, [19].
2. Identification
   a. The characteristic functions for \( \lambda_i|s_{it} \) and \( \sigma^2_i|s_{it} \) do not steadily disappear altogether into high-shrinkage processing method.
   b. Let \( \nu_{it} = \lambda_i + \epsilon_{it} \) be the composite error at time \( t \), the sequence \( \{\nu_{it}: t = 1,2,\ldots,T\} \) is almost certainly serially correlated, and definitely is if \( \{\epsilon_{it}\} \) is serially uncorrelated.
   c. With the panel setup in (1)m, with large \( N \) and sufficiently large \( T \), \( x_{it} \) includes interactions of variables with time periods dummies and general non-linear functions and interactions, so the model is quite flexible.
   d. For the CRE approach, each kind of covariates in \( x_{it} \) is separated out and the heterogeneous factors in \( \lambda_i \) are correlated with them.

Assumption 4.2 (Identification: (Un)balanced Panel) For all \( i \):

1. Identification
   a. \( s_{i0} \) is observed.
   b. \( (y_{i,0:T},x_{i,0:T},z_{i,0:T}) \) are observed.

2. Sequential Exogeneity (Conditional on the Unobserved Effects)
   a. \( E(y_{it}|s_{it},s_{it-1},\ldots,s_{i1},\lambda_i) = E(y_{it}|s_{it},\lambda_i) = \theta s_{it} + \lambda_i \) for \( i = 1,2,\ldots,N \).
   b. Sequential exogeneity is a middle ground between contemporaneous and strict exogeneity. It allows lagged dependent variables and other variables that change in reaction to past shocks.
   c. Because including contemporaneous exogeneity, standard kinds of endogeneity, where some elements of \( \lambda_i \) are correlated with \( \epsilon_{it} \), are ruled out such as measurement error, simultaneity, and time-varying omitted variables.
   d. Sequential exogeneity is less restrictive than strict exogeneity imposing restrictions on economic conditions.

3. Model Setup
   a. The term “correlated random effects” is used to denote situations where one models the relationship between \( \{\lambda_i\} \) and \( \{s_{it}\} \).
   b. The CRE approach allows to unify fixed and random effects estimation approaches.
   c. With the CRE approach, time-constant variables can be included within the system.

Statement 4.2.1 (Posterior Distributions) Under Assumptions (4.1) and (4.2), all posterior distributions in (28)–(33) hold and are estimable through MCMC algorithms and implementations.

Theorem 4.3 (Posterior Consistency: Correlated Random Coefficients) Given the DPD in (1):

1. Model: Assumptions (4.1) and (4.2).
2. Covariates: \( \left( y'_{i,0:T-1}, \gamma_{i,0:T-1}, z_{i,0:T}, z'_{i,0:T-1}, z_{i,0:T-1} \right) \) satisfy Assumptions (4.1) and (4.2).
3. Common Parameters:
   a. \( \nu_{it} \) is unknown and estimable.
   b. The domain of \( \sigma^2 \) is finite and estimable (homoskedastic case).
   c. \( \theta_0 \) is in the interior of \( \text{supp}(\Pi^0) \).
   d. For homoskedastic models, the domain of \( \sigma^2 \) is bounded by \( [\sigma^2, \hat{\sigma}^2] \) for some \( \sigma^2, \hat{\sigma}^2 > 0 \), where \( \sigma^2 \) (\( \hat{\sigma}^2 \)) is some small (large) positive number.

4. Stick-breaking Process: The base distribution of the MGLR prior in (19) is characterized by a multivariate normal distribution on \( \mu_F \) and an inverse Wishart distribution on \( \Omega_F \) with arbitrary scale parameters sufficiently small and arbitrary degrees of freedom close to zero. Thus, posterior distributions are strongly consistent at \( (\theta_0, l_0) \).

Statement 4.3.1 (Density Forecasts: Correlated Random Coefficients) Given conditions in Theorem 4.3, density forecasts converge to the oracle.

Theorem 4.4 (Density Forecasts: General Semiparametric Model) Given the DPD in Eq. (1):

1. Posterior Consistency: Conditions in Theorem 4.3.
2. Model:
   a. \( l_0 \) is bounded above by some positive numbers \( L \), with \( L > 0 \).
   b. The second moment of \( \varphi \) exists and is positive.
   c. If \( l \) is a conditional distribution, there exists \( q > 0 \) such that \( |q_0(s_{i0})| \geq q \) for all \( s_{i0} \in \mathcal{B} \), with \( q_0 \) denoting an induced measure.

3. Likelihood:
   a. \( \pi(y_{i,1:T} | \varphi, \theta_0, F) \) is continuous in \( \varphi \).
   b. \( 0 < \pi(y_{i,1:T} | \varphi, \theta_0, F) \leq L(F_1) \) for some \( L(F_1) > 0 \).
There exists \( q_1 > 0 \) such that for all \( \| \theta - \theta_0 \|_2 < q_1 \) and for some \( \{ B_y(F_i), B_y(F_i) \} > 0 \):

\[
a. \quad \int | \pi(\gamma_{i,t:T} | \varphi_i, \theta, F_i) - \pi(\gamma_{i,t:T} | \varphi_i, \theta_0, F_i) | d \varphi_i \leq B_y(F_i) \\
\| \theta - \theta_0 \|_2 \cdot \\
b. \quad \| \pi(\gamma_{i,T+m} | \varphi_i, \theta, F_i) - \pi(\gamma_{i,T+m} | \varphi_i, \theta_0, F_i) \|_1 \leq B_y(F_i) \\
\| \theta - \theta_0 \|_2 \cdot \\
Then, given individuals \( i \)'s density forecasts converge to the oracle in probability for all \( N \to \infty \) and \( \epsilon > 0 \):

\[
P(\| \gamma_{i,T+m}^{\text{cond}} - \gamma_{i,T+m}^{\text{oracle}} \|_1 < \epsilon | F_i) \to 1 \quad (37)
\]

**Empirical evidence**

**Data description and results**

The dataset contains 53 observable variables accounting for (potentially) covariates described through the vectors \( y_{i,t-1}, x_{it}, \) and \( z_{i,t-1} \). In this study, I split them in six groups: \( i \) lagged outcomes \( (y_{i,t-1}) \); \( ii \) individual-specific characteristics \( (y_{i,t-1}) \); \( iii \) economic indicators \( (z_{i,t-1}) \); \( iv \) healthcare and social factors \( (z_{i,t-1}) \); \( v \) exogenous variables \( (x_{it}) \); and \( vi \) lagged predictors concerning some spatial aspects based on the arbitrary delineation of the spatial units (e.g., county boundaries), the spatial distribution of health systems (e.g., health care expenditures), and the inherent mobility of county residents (e.g., physical activity). In this study, spatial units are dealt with the panel framework involved in the estimation, the added value per employee denotes the variable of interest accounted for evaluating cross-unit wage effects, and individual’s weight is used to estimate cross-unit excess weight across Italian regions. The latter has been computed according to the CDC's BMI = \( \frac{kg}{m^2} \), where \( kg \) is a person’s weight in kilograms and \( m^2 \) is its height in metres squared. It is so defined: underweight (below 18.5); normal (18.5 - 24.9); overweight (25.0 - 29.9); and obese (30.0 and above). In this study, I focus on the last two indicators.

The estimation sample is expressed in years and accounts for all the 21 Italian regions spanning the period 2006 – 2018. All data come from the Central Institute of Statistics (ISTAT) database according to the report on equitable and sustainable well-being (BES). Given the hierarchical structural conformation of the model and a sufficiently large number of years describing socio-economic and health issues, it is able to capture: \( i \) endogeneity issues; \( ii \) interdependency, commonality, and homogeneity; \( iii \) heterogeneous effects; and \( iv \) model misspecification problems.

According to the TROB procedure, in the first stage, I find that 39 best covariates better fit the data with posterior Inclusion Probabilities (PIPs) \( \geq r \) in (10) and \( \theta = 1 \). Thus, I obtain \( 2^{39} \) best model solutions \( \{ M_k \subset \mathcal{C} \} \). Because of the curse of dimensionality, I further shrink the data performing the second stage involved in the TROB procedure. In Table 1, I display the final 22 best covariates, obtaining \( 2^{22} \) best model solutions \( \{ M_k \subset \mathcal{C} \} \), with PIPs \( \geq \hat{r} \) in (11). All variables and descriptive statistics are displayed in Tables 4 and 5 on Appendix A, respectively.

Here, some considerations are in order. \( i \) Most model uncertainty and overfitting are avoided: indeed, the Conditional Posterior Sign (CPS) deals well with the sign certainty, taking values close to 1 (or 0) if a covariate in \( s_{ij} \) has a positive (or negative) effect on wage. \( ii \) Healthcare and social statistics matter as much as economic variables when evaluating cross-sectional wage effects (similar PIPs). \( iii \) Individual-specific factors and exogenous indicators need to be assessed when studying the effects of socio-economic factors on labour market outcomes (highly large PIPs). \( iv \) Cross-sectional individual factors are supposed to have homogeneous effects on wage. Contrarily, all non-homogeneous effects would be captured by \( \lambda_i \) via nonparametric Bayesian inference. \( iv \) The strictly exogenous factors \( (x_{it}) \) contain dummy variables to assess and then test the presence of structural breaks, evaluated through the Chow test, to investigate further interdependency among covariates and outcomes when performing density forecasts and dealing with (potential) spatial spillover effects. Indeed, labour market outcomes at a regional level are usually characterized by spatial dependence and then overweight may also exhibit spatial interdependence (see, for instance, [26] and [34]). \( v \) Finally, all predictors with PIPs \( \geq \hat{r} \) displayed in bold correspond to the ones to be accounted for the final solution.

Nevertheless, although the intense shrinkage in the parameter space, the final solution would still require more effort: indeed, there are 22 potential best (endogenous) predictors better fitting the data. Thus, I run the last stage involved in TROB procedure to obtain the final model solution better performing the data with higher log Bayes Factor. It consists of 16 final best covariates with \( \text{IBF} = 17.41 \) and so split: predictors \( (1, 2, 3, 5, 6, 8, 9, 11) \) for \( \gamma_{i,t-1} \); predictors \( (13, 14, 15, 17) \) for \( \gamma_{i,t-1} \); and predictors \( (18, 19, 20, 21) \) for \( \gamma_{i,t-1} \). All their lags are then involved as instruments in the estimating procedure to deal with endogeneity issues and functional forms of misspecification, with particular emphasis on the lagged outcomes \( (y_{it-1}) \) and predictors \( 1, 2, 5, 11 \) addressing (potential) causal effects and spatial dependence, respectively (Table 2). In the dynamic panel data regression, I also include three time-invariant effects \( (x_{1it}, x_{2it}, \) and

\(^4\) CDC stands for ‘Centers for Disease Control and Prevention’ (https://www.cdc.gov’cdc.gov’).
tested through the Chow test, denoting the presence of (potential) structural breaks in: 2007, due to the global financial crisis; 2011, due to fiscal consolidations in post-financial crisis; and 2014, due to recovery funds. Moreover, because of possible significant difference between males and females affecting outcomes, I run three different models: (i) Model 1 accounting for all individuals’; (ii) Model 2 focusing on male individuals; and (iii) Model 3 accounting for female individuals. Gender disparities have been tested through F-test by comparing the variances of the two samples. More precisely, I test whether there are no differences among the two variances for males and females (H0 : σ²M = σ²F) versus the two-sided alternative (H0 : σ²M ≠ σ²F). Overall, the coefficients differ between males and females with: F = 61.27; p-value = 0.001; and ratio of explained variances equals to δ²T = 79.43%, where T stands for 'Total variance'.

5 For instance, the ‘Recovery Assistance for Cohesion and the Territories of Europe’ (REACT-EU) programme for 2014-2020.

The DPD estimates are displayed in Table 2. Concerning Model 1, six findings are addressed. (i) Health status and social factors strongly affect wage rates and are essential in determining a stable and well-functioning economy, high-quality jobs, and then the stability of a country. The negative relationship between health sector and wage (predictors 1, 2, 3, 5, 9) highlights the positive impact that the former has on the economic performance in the national economy through the jobs it generates and the purchase of goods and services. (ii) Healthcare factors would reduce social exclusion at the local level (significance and positive magnitude for predictor 8) due to their impact on employment, working conditions, and household income. Thus, the health sector can increase economic status acting as a key sector for driving forward the implementation of local and national goals for sustainable development. For instance, health and social work activities have constituted around 10% of total employment in 2015 according to the Organisation for Economic Co-operation and Development (OECD) countries. (iii) Accounting for economic and individual-specific factors, the availability
Table 2 Dynamic regression results

| Idx | Variables | Model 1 - total | Model 2 - male | Model 3 - female |
|-----|-----------|----------------|---------------|-----------------|
| Healthcare and social statistics | | | | |
| 1   | Overobe   | −1.84 **        | −0.98 **      | −1.23 **        |
| 2   | Obe       | −1.92 ***       | −1.16 ***     | −1.21 ***       |
| 3   | Smoke     | −1.33 **        | −0.94 *       | −1.35 ***       |
| 5   | Sed       | −0.57 **        | −0.68 **      | −0.81 ***       |
| 6   | Food      | −1.25 ***       | −1.15 ***     | −1.32 ***       |
| 8   | Socpar    | 0.52 *          | 0.32 *        | 0.25 **         |
| 9   | Rop       | −0.57 **        | −0.44 *       | −0.46 **        |
| 11  | Hexp      | 1.35 ***        | 1.46 ***      | 1.49 ***        |
| Economic status | | | | |
| 13  | Degree    | 0.92 ***        | 1.11 ***      | 0.61 ***        |
| 14  | Employ    | 3.77 ***        | 3.77 ***      | 3.72 ***        |
| 15  | Fterm     | −1.94 ***       | −1.68 ***     | −1.74 ***       |
| 17  | Exprd     | 9.56 ***        | 8.91 ***      | 8.53 ***        |
| Individual-specific factors | | | | |
| 18  | Nsw       | −2.42 ***       | −2.26 ***     | −2.31 ***       |
| 19  | Family    | 0.78 ***        | 0.73 **       | 0.75 **         |
| 20  | Fime      | 0.44 ***        | 0.37 **       | 0.37 **         |
| 21  | Pjob      | −2.75 ***       | −2.69 ***     | −2.72 ***       |
| Exogenous factors | | | | |
| – x1t | −1.09 *** | −1.14 *** | −1.33 *** |
| – x2t | 1.55 ***  | 1.66 ***     | 1.58 ***     |
| – x3t | 2.18 ***  | 2.22 ***     | 2.17 ***     |
| Lagged predictors - spatial dependence | | | | |
| – L. overobe | −1.89 ** | −1.74 ***     | −1.67 ***     |
| – L.obe     | −1.03 *** | −1.75 ***     | −1.79 ***     |
| – L. sed    | −0.67 **  | −0.73 **      | −0.76 ***     |
| – L. hexp   | 1.59 ***  | 1.63 ***      | 1.65 ***      |
| Lagged outcomes - dependent variable | | | | |
| – L. ave    | 0.54 ***  | 0.51 ***      | 0.55 ***      |

The Table is so split: the first two columns denote the predictor number and label; and the last three columns display the estimates and the standard errors (in parenthesis). Here, the variable of interest is the added value per employee and L. stands for the first order lag operator concerning the outcomes and some endogenous predictors to deal with causal relationships and spatial dependence, respectively. The significant codes are: *** significance at 1%; ** significance at 5%; and * significance at 10%

of good jobs lies at the heart of inclusive and sustainable growth eliminating poverty. More precisely, good working conditions would provide: decent pay ensuring at least a minimum or living wage (positive magnitude for predictors 19, 20); employee benefits, such as maternity and paternity leave, ensuring part- and full-time workers to receive similar benefits (negative magnitude for predictor 21); minimal use of temporary contracts (negative magnitude for predictor 15); safe working conditions (negative magnitude for predictor 18); and opportunities for progression and career development (positive magnitude for predictors 13, 14, 17). (iv) Overweight (predictor 1), obesity (predictor 2), and sedentary rate (predictor 5) are likely simultaneously determined in the estimating procedure to deal with spatial dependence and produce unbiased, consistent, and efficient estimators. Here, simultaneity arises when a predictor is endogenous to the system and then is likely correlated with the error term. (v) According to the previous point, let the PIPs and the degree of relationship between wage effects and excess weight (predictors 1 and 2) be very close, it highlights how environmental and policy issues, including interventions to change the local environment to create opportunities for physical activity, would positively affect the economic dynamics and then its performance in the national economy. Moreover, the % of obese individuals would tend to matter slightly more than the % of overweight ones, emphasizing the negative impact that an excessive weight has not only on work activities but also on the health system (e.g., higher magnitude of excess lifetime medical care costs associated with an increasing degree of obesity). These findings find confirmation with related previous studies such as [70] among others. (vi) Finally, concerning time-fixed effects, policy programs (x1), positively affect outcomes with larger magnitude than post-crisis fiscal measures (x2). Conversely, negative effects on outcomes are displayed during the Great Recession (x3).

According to gender (Models 2 and 3), three main results are in order. (i) First, employment penalties persist for males and even more for females. The undervaluing of women’s social care and health jobs relate to wider norms and attitudes in society, where women’s skills are often less visible than men’s ones and harder to be quantified, mainly when linked to productivity. Generally, women are stereotyped as being naturally good at the job and are thought to be prepared to trade lower pay for job reward. Despite the continuing reduction in the disparity between male and female employment rates, large gender differences still remain in Italy. These findings recognize the results found in previous studies such as [11, 24], and [30]. (ii) Second, healthcare and social factors negatively affect employment opportunities and wage rates for males and slightly more for females. For instance, negative impacts on working conditions, e.g., due to smoking, alcohol use, and inappropriate use of food, would matter more for females. (iii) Third, another important factor to be accounted for understanding the pay levels of health and social care workers is that many more women than men work part-time hours in employment. Indeed, individual-specific differences highlight that discrimination still accounts for a significant amount of the gender pay gap, referring to the relative differences in average gross hourly earnings of women and men. However, the gender
Table 3  Diagnostic panel tests

| Test statistic | Model 1 - total | Model 2 - male | Model 3 - female |
|----------------|----------------|---------------|-----------------|
| Significance and Robustness | \( R^2 \) | 0.89 | 0.85 | 0.86 |
|  | \( Q_{st} \) | 255.03 (0.00) | 157.02 (0.00) | 197.01 (0.00) |
|  | \( Q_d \) | 1.78 (0.19) | 2.86 (0.11) | 0.99 (0.32) |
|  | \( Q_{LR} \) | 0.02 (0.89) | 0.07 (0.79) | 0.04 (0.84) |
| Spatial dependence | Robust LM on overobe | 9.97 (0.00) | 11.30 (0.00) | 11.64 (0.00) |
|  | Robust LM on obe | 10.14 (0.00) | 12.07 (0.00) | 12.74 (0.00) |
|  | Robust LM on sed | 5.44 (0.03) | 5.83 (0.02) | 6.03 (0.00) |
|  | Robust LM on hexp | 12.87 (0.00) | 13.12 (0.00) | 13.46 (0.00) |
| Panel unit root tests | \( U_H \) | \(-4.16 (0.13)\) | \(-4.35 (0.21)\) | \(-4.68 (0.23)\) |
|  | \( U_I \) | \(-3.10 (0.09)\) | \(-4.16 (0.02)\) | \(-3.71 (0.03)\) |
|  | \( U_M \) | \(-3.38 (0.06)\) | \(-3.23 (0.08)\) | \(-3.42 (0.07)\) |
| Time-fixed effects | \( F_{s1} \) | 31.43 (0.00) | 22.86 (0.00) | 25.53 (0.00) |
|  | \( F_{s2} \) | 16.04 (0.00) | 17.73 (0.00) | 15.59 (0.00) |
|  | \( F_{s3} \) | 15.22 (0.00) | 4.92 (0.00) | 5.74 (0.00) |

The Table refers to the results of diagnostic tests for significance and robustness in the estimating procedure, autocorrelation problems, panel unit root tests, spatial dependence, and structural breaks.

Pay gap needs to be evaluated in either the economic status or healthcare and social context to understand the position of women in society. More precisely, in specific occupation rates of pay, women are likely to work in jobs that undervalue their skills and training, and to be numerically lower than in those where men predominate. These results are consistent according to the structural indicators drawn up by the European Union to monitor the implementation of the European Strategy for Jobs and Growth (see, for instance, the European Commission in the Roadmap for Gender Equality, 2006 – 2010). According to time-fixed effects, the results are similar to the ones found in Model 1, with a slightly larger (lower) magnitude for females according to negative (positive) effects on outcomes.

In Table 3, I display the main diagnostic tests dealing with significance and robustness of the estimating procedure, autocorrelation problems, panel unit root tests, spatial dependence, and structural breaks. Here, some considerations are in order. (i) The dynamic panel regression shows a \( R^2 \geq 89\% \) and then is robust evaluating most of the explained variability of the outcomes of interest, with \( R^2 \) referring to the adjusted \( R^2 \). (ii) The Sargan-Hansen’s test for over-identification (\( Q_{st} \)) highlights the performance and usefulness of the DPD to deal with endogeneity issues and functional forms of misspecification ([68] and [37]). Moreover, it is also able to address variable selection problems avoiding model uncertainty and overfitting (for instance, the CPSs close to 1 or 0 in Table 1). (iii) There are no serial correlations in idiosyncratic errors according to the first order lag operator. The test statistic refers to [6] (\( Q_d \)). (iv) The Multivariate Ljung-Box test (\( Q_{LR} \)) shows no serial correlation among residuals over time and then the time-series results are robust and valid ([49]). (v) The highly large \( R^2 \) and Sargan’s test statistics would also highlight the importance to account for heterogeneous individual characteristics and socio-economic factors when studying the effects of healthcare status, overweightness, and labour market outcomes. These relationships are emphasized by testing the spatial dependence through the [3]’s robust Lagrange Multiplier (LM) statistics on the autoregressive coefficients \( \beta \) and \( \gamma \) in (1). The statistics are distributed as a \( \chi^2 \) with one degree of freedom, where the null stands for no spatial autocorrelation. The results are robust to the presence of spatially lagged correlation in the variable of interest (ave) when testing for spatial correlation in excess weight (overobe and obe), and healthcare statistics (sed and hexp). (vi) Concerning panel unit root tests, I use three different test statistics for identifying unit root processes and trend-stationarity for every \( AR(1) \) time series involved in the system as instrument: (1) [35] \( U_H \) that extends the [43]’s test to panel data with i.i.d. heteroskedastic and serially correlated disturbance terms across \( i \) and over \( t \) by proposing a residual-based Lagrange multiplier test; (2) [39] \( U_I \) that suggests standardized t-bar test statistic based on the Augmented Dickey-Fuller statistics averaged across units; and (3) [52] \( U_M \) that proposes a comparative study of unit root tests with panel data. In the first study, stationarity around a deterministic trend (or level) is the null hypothesis against the alternative of a unit root. Conversely, in the last two studies, the unit root is the null hypothesis to be tested against the alternative of trend-stationarity. In all these test statistics, either the critical values or the moments of the asymptotic distributions have been evaluated using MCMC algorithms. The results confirm no unit root in the lagged time series and then likely ensure trend-stationarity. (vii) Finally, potential structural breaks matter for every estimation model affecting the outcomes. They are dealt with the three time-fixed effects included in the estimating procedure through \( F \)-statistics (\( F \)).

Density forecasts and policy-relevant strategies

In Fig. 1, I draw conditional density forecasts through accurate MCMC algorithms and implementations. More precisely, posterior distributions are computed according to Eqs. (28)–(30) for \( \hat{\theta}_t | \tilde{\theta}_{t-1} \) and \( \tilde{y}_{t+\tau} | \tilde{\theta}_{t-1} \). (34)–(35) for moment distributions in \( \lambda_t | \tilde{\theta}_{t+\tau} \) given initial values (\( s_0 \),
Obesity and labour market outcomes in Italy: a dynamic panel data evidence with correlated random…

Density forecasts are performed running 100,000 iterations per each random start, spanning the period 2006 to 2021. The optimal number of draws has been chosen according to the IBF: indeed, with 100 and 1000 iterations, I obtain highly lower IBF (4.89 and 5.34, respectively). The associated computational costs are minimized, ensuring consistent posterior estimates and dimension reduction. Density forecast combinations correspond to the projections of every subsequence drawn in the sample. In this context, I divide the sample into three groups to better highlight the strengths and limits of policy strategies: (1) Northern Italy; (2) Central Italy; and (3) Southern Italy. The yellow and red curves denote the 95% confidence bands, and the blue and purple curves denote the conditional\(^6\) and unconditional\(^7\) projection of outcomes \(\hat{y}_{i,t+\tau}\) for each time period \(t\), respectively. Here, the outcomes absorb the conditional forecasts computed for a time frame of 3 years in order to also address potential findings concerning the impact of the ongoing pandemic crisis on the global economy. The natural conjugate prior refers to the three subsamples according to the structural breaks involved in the panel regression: 2007; 2011; and 2014.

The results emphasize the findings obtained in Table 2. (i) Conditional projections lie in the confidence interval; conversely, unconditional projections tend to diverge over time. Thus, when studying time-varying healthcare and socio-economic factors, structural time-fixed effects \((x_{it})\) and correlated individual-random effects \((\lambda_i)\) have to be accounted for. (ii) Density forecasts in group 1 (top-plot) tend to show a similar evolution, but with a larger spillover effects to the ones in group 2 (middle-plot) because of not directly observed regional and European country interdependencies (e.g., similar dynamics, co-movements, economic structure). (iii) Density forecasts in group 3 (bottom-plot) tend to display larger divergences, in terms of spillovers, because of potential functional forms of misspecification (e.g., strong inter-linkages across regions in Northern Italy driving the transmission of unexpected shocks on labour market outcomes). (iv) Most positive projections in groups 1 and 2 highlight the accuracy of the methodology by dealing with endogeneity.

---

\(^6\) Generally, the conditional projection in density forecasts is the one that the model would have obtained over the same period conditionally on the actual path of unexpected dynamics for that period.

\(^7\) Generally, the unconditional projection in density forecasts is the one that the model would obtain for output growth for that period only on the basis of historical information, and it is consistent with a model-based forecast path for the other variables.
issues. For instance, regions in the Northern Italy tend to be net sender because of consistent interdependencies among them and Northern Europe, and then Central Italy regions because of stringent economic-institutional linkages with the former. (v) The spillover effects show a sudden change in their dynamics according to the three structural breaks evaluated in the panel data regression. Thus, it highlights the presence of persistent (unobserved) shocks over time affecting time-varying predictors and then outcomes. (vi) Accounting for the ongoing pandemic crisis (conditional projections in 2020 – 2021), a sufficiently stronger recovery seems to matter across the regions in Northern Italy. Conversely, slower socio-economic improvements would affect the regions in Central Italy and, even more, the southwards. Thus, the need for forecasters and policymakers to deal with cross-sectional individual-specific heterogeneity when investigating dynamic feedback between socio-economic effects and labour market outcomes.

From a policy perspective, the above empirical results can be so summarized. (i) Appropriate prevention policy strategies in terms of higher healthcare and social statistics need to be improved, mainly in accordance with the current pandemic disease. Indeed, solid and coordinated improvements in healthier and safer workplace would increase employment opportunities and wage effects, achieving higher labour productivity and (possibly) lessening social disparities in gender. (ii) Better and more active workforce should be designed to deal with key behavioural and individual-specific risk factors, such as obesity, social participation, non-regular employments, and leisure satisfaction, negatively affecting employment opportunities and wage rates for males and slightly more for females. (iii) Since 2011, a substantial number of labour market and social policy reforms has been achieved, improving the flexibility of job creation and extending social assistance and social protection coverage. In 2014, further measures have been enacted under the so-called Jobs Act reform (‘Renzi government’). The main goal of the Jobs Act reform has been twofold. First, it has aimed to encourage jobs growth by making labour input cheaper. Second, the reform has aimed to make contracts more appealing to employers by the introduction of ‘increasing-protection employment contracts’. These refer to new permanent open-ended contracts, with lower dismissal costs and reduced legal uncertainty compared to the previous ones to promote a combination of labour market flexibility and security for workers. (iv) Recently, in 2018, a partial reversal of the Jobs Act provisions has been carried out by the Conte government, the so-called ‘Dignity Decree’, by restricting the use of fixed-term contracts and increasing the costs of dismissals. According to social policies, a combination of vocational education and employment measures has been also revised under the so-called ‘Good School reform’, by including initiatives to simplify short-term contracts and then align them better with labour market needs. Nevertheless, besides the weak labour market recovery in terms of job creation, the Italian labour market is still characterised by increasing instability of employment and use of fixed-term contracts.

From a global perspective, related considerations are in order. (i) European Union Structural Funds have supported the setting up the necessary infrastructure in terms of services and operators. More precisely, the ‘Stability Law’ in 2016 have introduced the ‘Support for Active Inclusion’ (SIA) to address the way for the new ‘Inclusion Income’ implemented in December 2017. However, due to the lack of resources, the new measure has been not sufficient to cover all individuals living in absolute poverty, and the monetary transfer to families follows to be still very small. (ii) Concerning the COVID-19 pandemic crisis, confirmed in Italy on 31 January 2020, it has rapidly expanded across regions in Northern Italy and then the southwards, although the impact at the local level has been heterogeneous (see, for instance, Figure 1). On 22 April 2020, Italy has been the third country in the world by the number of reported COVID-19 infection cases, and the second by the number of deaths among the infected patients (see, for instance, the ‘Center for Systems Science and Engineering’). The progressive tightening of containment measures has affected more and more larger segments of the Italian economy, mainly with regards to retail trade businesses with the exception of those related to the sale of food and other basic necessities. Since the beginning of March 2020, the OECD has designed relevant policy measures aimed at improving the healthcare system and providing economic support to households, workers, and businesses. Then, the second set of strategies has been also developed supporting employees, self-employed, and businesses. (iii) Finally, according to the European Union’s regional policies (2018) and the post-2020 Cohesion policy reform, several ‘place-based’ programs have been addressed for the redistribution of wealth across regions and countries to stimulate growth in the areas lagging in development. In that context, three main objectives can be summarized. First, public finances and current expenditure should be controlled but allowing development policies to be implemented in backward areas with the greatest potential. Second, an adequate system of fiscal compensation has to be included to offset the structural competitive disadvantages to which the Southern Italy regions are exposed, mainly within the Eurozone, with a view to progressively overcoming them. Third, a rebalancing of the current geopolitical configuration to aim for cooperation and development policies for the Southern area needs to be accounted for.

Concluding remarks

This study improves the existing literature when investigating the effects of obesity, socio-economic variables, and individual-specific factors on work productivity across Italian regions. A dynamic panel data with correlated random effects is used to address variable selection problems in high-dimensional time-varying data. Semiparametric Bayesian
inference is involved in the estimating procedure to jointly deal with incidental parameters, endogeneity issues, and functional forms of misspecification. The main thrust of the proposed methodology is the use of conjugate hierarchical informative (proper) priors to discover the most probable subset of potential predictors affecting outcomes to capture different and/or not directly observed dynamics and interconnections among lagged endogenous factors. Posterior distributions for conditional density forecasts are obtained by MCMC algorithms and implementations. Cross-sectional Heterogeneity is modelled nonparametrically allowing for correlation between heterogeneous parameters and initial conditions as well as individual-specific regressors.

An empirical example describes the estimating procedure and forecasting performance. More precisely, the estimation sample accounts for all the 21 Italian regions covering the period 2006 – 2018. All data come from the Central Institute of Statistics database according to the report on equitable and sustainable well-being. Conditional density forecasts are then performed spanning the period 2006 to 2021 to also address potential findings concerning the impact of the ongoing pandemic crisis on the global economy. In that context, the sample is divided in three groups to better highlight the strengths and limits of policy strategies: (1) Northern Italy; (2) Central Italy; and (3) Southern Italy. Moreover, natural conjugate priors are involved in the procedure to also deal with three permanent time-invariant effects affecting panel data: 2007, due to the global financial crisis; 2011, due to fiscal consolidations in post-financial crisis; and 2014, due to recovery fund.

### Data collection

Tables 4 and 5 display the covariates and their descriptive statistics involved in the analysis according to the second stage of the TROB procedure, respectively. They refer to the subset $E \in S$ [Eq. (11)] denoting all the best submodel solutions.

**Table 4** Data source – subset $E$

| Idx. | Variable                  | Source                                      |
|------|---------------------------|---------------------------------------------|
|      | Healthcare and social statistics |                                             |
| 1    | Overweight                | ISTAT: Aspects of Daily Life Survey        |
| 2    | Obesity                   | ISTAT: Aspects of Daily Life Survey        |
| 3    | Consumption of tobacco    | ISTAT: Aspects of Daily Life Survey        |
| 4    | Consumption of alcohol    | ISTAT: Aspects of Daily Life Survey        |
| 5    | Sedentary rate            | ISTAT: Aspects of Daily Life Survey        |
| 6    | Appropriate use of food   | ISTAT: Aspects of Daily Life Survey        |
| 7    | Cultural interests        | ISTAT: Aspects of Daily Life Survey        |
| 8    | Social participation      | ISTAT: Aspects of Daily Life Survey        |
| 9    | Risk of poverty           | ISTAT: ‘EU-SILC’ Survey                    |
| 10   | Public transport satisfaction | ISTAT: Aspects of Daily Life Survey    |
| 11   | Health care expenditure   | ISTAT: Aspects of Daily Life Survey        |
|      | Economic status           |                                             |
| 12   | High school diploma       | ISTAT: Labour Force Survey                 |
| 13   | Graduates/other qualifications | ISTAT: Labour Force Survey          |
| 14   | Employment rate           | ISTAT: Labour Force Survey                 |
| 15   | Fixed-term contract (= 5 years) | ISTAT: Labour Force Survey  |
| 16   | Weighted income per capita | ISTAT: National Accounts                 |
| 17   | Expenditure on R &D        | R &D in Enterprises Survey                 |
|      | Individual-specific factors |                                             |
| 18   | Neither studying nor working | ISTAT: Labour Force Survey               |
| 19   | Family relationship satisfaction | ISTAT: Aspects of Daily Life Survey |
| 20   | Free time satisfaction     | ISTAT: Aspects of Daily Life Survey        |
| 21   | Poorly-paid job            | ISTAT: Labour Force Survey                 |
| 22   | Non-regular employment     | ISTAT: National Accounting                |
| 23   | Added value per employee   | Small- and Medium-sized Enterprises (SMEs) Survey |

Here, ‘EU-SILC’ stands for ‘European Union-Income and Living Conditions’ Survey (https://www.istat.it/en/archivio/216934#eu-silc.istat). All data come from ISTAT database (https://www.istat.it/en/archive#documents_archive)
Funding  No funding was used for this study.

Declarations

Conflict of interest  The author declares no conflict of interest.

Ethical approval  This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Adams, P., Hurd, M.D., McFadden, D., Merrill, A., Ribeiro, T.: Healthy, wealthy, and wise? tests for direct causal paths between health and socioeconomic status. J. Econom. 112(1), 3–56 (2003)
2. Agerström, J., Rooth, D.O.: The role of automatic obesity stereotypes in real hiring discrimination. J. Appl. Psychol. 96(4), 790–805 (2011)
3. Anderson, T.W., Hsiao, C.: Estimation of dynamic models with error components. J. Am. Stat. Assoc. 76(375), 598–606 (1981)
4. Angrisani, L., Santonicola, A., Iovino, P., Vitiello, A., Higa, K., Himpens, J., Buchwald, H., Scopinaro, N.: Ifso worldwide survey 2016: Primary, endoluminal, and revisional procedures. Obes. Surg. 28, 3783–3794 (2018)
5. Anselin, L., Bera, A.K., Florax, R., Yoon, M.J.: Simple diagnostic tests for spatial dependence. Reg. Sci. Urban Econ. 26(1), 77–104 (1996). https:// doi.org/ 10. 1016/ 0166- 0462(95) 02111-6
6. Arellano, M.: Panel data econometrics. Oxford University Press, New York (2003)
7. Arellano, M., Bond, S.: Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. Rev. Econ. Stud. 58(2), 277–297 (1991)
8. Arellano, M., Honore, B.: Panel data models: some recent developments. In: Heckman, J., Leamer, E. (eds.) Handbook of Econometrics, vol. 5, pp. 3229–3296. Elsevier (2001)
9. Barnay, T.: Health, work and working conditions: a review of the European economic literature. Eur. J. Health Econ. 17, 693–709 (2015)
10. Bateman, I.J., Jones, A.P., Lovett, A.A., Lake, I.R., Day, B.H.: Applying geographical information systems (gis) to environmental and resource economics. Environ. Resour. Econ. 22(1–2), 219–269 (2002)
11. Baum, C.L., Ford, W.F.: The wage effects of obesity: a longitudinal study. Health Econ. 13, 885–899 (2004)
61. Puhl, R.M., Brownell, K.: Bias, discrimination, and obesity. Obes. Res. 9, 788–805 (2001)
62. Rashad, I., Grossman, M., Chou, S.: The super size of America: an economic estimation of body mass index and obesity in adults. East. Econ. J. 32(1), 133–148 (2006)
63. Register, C.A., Williams, D.R.: Wage effects of obesity among young workers. Soc. Sci. Q. 71(1), 130–141 (1990)
64. Rodriguez, A., Dunson, D.B.: Nonparametric Bayesian models through probit stick-breaking processes. Bayesian Anal. 6(1), 145–178 (2011)
65. Roehling, M.V., Roehling, P.V., Pichler, S.: The relationship between body weight and perceived weight-related employment discrimination: the role of sex and race. J. Vocat. Behav. 71(2), 300–318 (2007)
66. Rooth, D.-O.: Obesity, attractiveness, and differential treatment in hiring: a field experiment. J. Hum. Resour. 44, 710–735 (2009)
67. Rosen, H.S., Wu, S.: Portfolio choice and health status. J. Financ. Econ 72(3), 457–484 (2004)
68. Sargan, J.D.: The estimation of economic relationships using instrumental variables. Econometrica 26, 393–415 (1958)
69. Sobal, J.: Sociological analysis of the stigmatisation of obesity. In: Germov, J., Williams, L. (eds.) A sociology of food and nutrition: the social appetite, 2nd edn. Wiley, Melbourne (2004)
70. Thompson, D., Edelsberg, J., Colditz, G.A., Bird, A.P., Oster, G.: Lifetime health and economic consequences of obesity. Arch. Intern. Med. 159, 2177–2183 (1999)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.