A Two-step System for Hierarchical Bayesian Dynamic Panel Data to deal with Endogeneity Issues, Structural Model Uncertainty, and Causal Relationship

Pacifico, Antonio

LUISS Guido Carli University

2020

A Two-step System for Hierarchical Bayesian Dynamic Panel Data to deal with Endogeneity Issues, Structural Model Uncertainty, and Causal Relationship

Antonio Pacifico*

Abstract

The paper develops a computational method implementing a standard Dynamic Panel Data model with Generalized Method of Moment (GMM) estimators to deal with endogeneity issues, structural model uncertainty, and causal relationship in large and long panel databases. The methodology takes the name of Two-step System Dynamic Panel Data, that combines a first-step Bayesian procedure for selecting the only potential predictors in a static linear regression model with a frequentist second-step procedure for estimating the parameters of a dynamic linear panel data model. An empirical example to the effects of obesity, socioeconomic variables, and individual-specific factors on labour market outcomes among Italian regions is performed. Potential prevention policies and strategies to address key behavioural and diseases risk factors affecting labour market outcomes and social environment are also discussed.

JEL classification: A1; C01; E02; H3; N01; O4.

Keywords: Bayesian Model Averaging; Dynamic Panel Data; Granger Causality; Labour Market Outcomes; Obesity.

*Corresponding author: Antonio Pacifico, Postdoc in Applied Statistics & Econometrics, LUISS Guido Carli University, CEFOP-LUISS (Rome), CSS Scientific Institute (Tuscany). Email: antonio.pacifico86@gmail.com or apacifico@luiss.it. ORCID:https://orcid.org/0000-0003-0163-4956

1 Introduction

Dynamic Panel Data (DPD) regressions are basically subject of estimation bias over time. Since the lagged dependent variable Y_{t-p} or the lagged explanatory variables X_{t-p} could be endogenous, with t and p denoting generic time and lag periods, their presence may cause correlation with the error term u_t. In addition, when studying and investigating multicountry economic interactions and policy implications in a context of large dynamic panels, endogeneity issues – because of unobserved heterogeneity and/or omitted factors – and structural model uncertainty – where one or more parameters are posited as the source of model misspecification problems – can occur among study units (see, e.g., Pacifico (2019c,a) and Pacifico (2020a)). Thus, with dynamic and endogenous variables, the use of the Generalized Least Squares (GLS) or the Fixed Effects (FE) estimators would lead to inconsistent estimates (see, e.g., Baltagi (1995)). Furthermore, when both N and T are large, Granger-causality relationship needs to be tested in panel setups. The basic idea is that if past values of X are significant predictors of the current value of Y even when past values of Y have been included in the model, then X exerts a causal influence on (or Granger-causes) Y (see, e.g., Dumitrescu and Hurlin (2012), Harris and Tzavalis (1999), Im et al. (2003), Levin et al. (2002), and Pesaran (2007)).

The ability of fixed-effects technique – and first differencing as well – to remove endogeneity issues has been largely proved in the context of a DPD model. Nevertheless, a serious difficulty occurs because the demeaning process which subtracts the individual's mean value of the outcome Y and each covariate X from the respective variable creates a correlation between predictor and error, particularly in the small T and large N context, where T and N denote generic time periods and individual units[1]. The resulting correlation creates a bias in the estimated coefficients of the lagged outcome which is not mitigated by increasing N. By including additional factors, it does not remove the bias: indeed, if the predictors are correlated with the lagged outcome to some degree, their coefficients may follow to be seriously biased. The same problem arises from the (one-way) random effects model. The causal component enters every value of Y by assumption, so that the lagged outcome cannot be independent of the composite error process[2].

The methodological contribution of this article faces up to these limits and overtakes them by modeling and implementing a standard Dynamic Panel Data with GMM estimators to jointly deal

[1] See, for instance, Nickell (1981).

[2] The same problem occurs with the first difference transformation: indeed, even if it removes both constant intercepts and individual effects, there is still correlation between the differenced outcome and the disturbance process, which is now a first-order Moving Average (MA(1)) process.

with endogeneity issues, functional forms of misspecification, and causal relationship in large and long panel databases. The methodology consists of a two-step approach – labelled Two-Step System Dynamic Panel Data (TSDPD) procedure – that combines a first-step Bayesian procedure for selecting the only (potential) predictors in a static linear regression model with a frequentist second-step procedure for estimating the parameters of a dynamic linear panel data model.

The first step builds on Pacifico (2020b), who develops a Robust Open Bayesian (ROB) procedure – entailing two stages – for implementing Bayesian Model Selection (BMS) and Bayesian Model Averaging (BMA) in multiple linear regression models when accounting for dynamics of the economy in either time-invariant moderate data or time-varying high dimensional multivariate data. In this study, I apply the ROB procedure by performing the implicit fully enumerated Markov Chain Monte Carlo (MC^F) integration[3] on a set of cross-sectional data with time-invariant factors in order to find a pool of predictors with highly strong explanatory powers on the outcomes. In this way, I will be able to simultaneously move through the model and the parameter space and thus obtain a reduced set containing $best$ potential model solutions (or $best$ combination of predictors) that mainly explain and thus fit the data. Then, a further shrinkage is conducted in order to obtain a smallest final subset of $top\ best$ submodels containing the only $significant$ solutions. Finally, the submodel with higher Bayes Factor (BF) will be the final solution containing a subset of predictors having higher $significant$ overall F value and sufficiently $strong$ adjusted-R^2 ($\bar{R}^2$) measure. Here, $best$ and $top\ best$ stand for the model providing the most accurate predictive performance over all candidate models and submodels, respectively, $significant$ stands for models having statistically significant predictive capability, and $strong$ refers to $\bar{R}^2$ value equal to or bigger than 30%.

Given the final $top\ best$ sample, the second step entails the construction of a DPD model by including all available lags of the outcomes and predictors as instruments to obtain consistent and unbiased estimates. More precisely, I build on Arellano and Bond (1991), which popularize the work of Holtz-Eakin et al. (1988) based on the notion that a simple instrumental variable approach – e.g., by adding one or more lagged dependent variables to allow for the modeling of a partial adjustment mechanism[4] – does not exploit all of the information available in the sample. By doing so in a Generalized Method of Moments (GMM) context, one may construct more efficient estimates of the DPD model. A key aspect of the Arellano and Bond (1991)'s strategy is the assumption that the necessary instruments are $internal$ or based on lagged values of the in-

[3]See, for instance, Pacifico (2020b).
[4]See, for instance, Anderson and Hsiao (1981).

strumented variables. The GMM estimators allow the inclusion of *external* instruments as well. Thus, I follow the underlying logic, but with some novelty. More precisely, the *external* instruments are used to take into account all the available lags of the time-varying variables (either X_t or Y_t) and thus potential causal interactions. In this way, the model is able to deal with endogeneity issues because of omitted variables or unobserved heterogeneity. Moreover, a correlated random effects approach is used in which the unobserved individual heterogeneities are treated as random variables that are possibly correlated with some of the predictors within the system. In this way, possible biases in the estimated coefficients of lagged outcomes will be avoided as well.

The methodology proposed in this paper cosists in four main contributions: (i) use correlated random effects approach to address and then avoid (potential) correlations between lagged outcomes; (ii) avoid endogeneity issues when studying dynamic panel data; (iii) deal with variable selection problems to select the best combination of predictors affecting the outcomes (such as overfitting[5], model uncertainty[6], and choice and specification of prior distributions); and (iv) use external instruments to identify and thus investigate (potential) causal links among covariates and variables of interest.

The application and empirical analysis aim focus on the relationship between high body weight (obesity) and labour market outcomes across Italian regions, by including a set of potential predetermined variables[7] (e.g., lagged values of the variables of interest), endogenous variables (e.g., socioeconomic factors varying over time and thus possibly correlated with contemporaneous errors), and heterogeneous individual-specific factors possibly correlated with some variables within the system. The time period spans the years between $2007-2017$ in order to cover a sufficiently large sample to address possible causal relationships between obesity, wages, and labour productivity. Furthermore, the empirical strategy is also able to investigate and thus design (potential) prevention policies and strategies to address key behavioural and diseases risk factors affecting labour market outcomes and social environment.

The outline of this paper is as follows. Section 2 discusses the econometric methodology describing in depth the two involved strategies. Section 3 presents a background literature and re-

[5]Overfitting and thus overestimation of effect size arise when variable selection procedure involves making inference on more complex models since they will always provide a somewhat better fit to the data than simpler models, where the 'complexity' stands for the number of unknown parameters. See, for instance, Pacifico (2020b).

[6]Overall, model uncertainty occurs when dealing with Bayesian inference and standard variable selection procedure. It arises when – given a set of all possible candidate covariates – a subset of potential covariates better explaining and thus fitting the data is obtained by conditioning on a single model and, then, making inferences as if the selected model has been the true model. See, for instance, Miller (1984), Breiman (1992, 1995), and Breiman and Spector (1992).

[7]In econometrics, predetermined variables denote covariates uncorrelated with contemporaneous errors, but not for their past and future values.

lated works on the relationship between high body weight and labour market outcomes. Section 4 illustrates the empirical analysis across regions in Italy, with a particular emphasis on possible causal links between obesity and adverse labour market outcomes for designing effective public policy. Section 5 contains some concluding remarks.

2 The Econometric Methodology

2.1 Bayesian Framework - First Step

According to Pacifico (2020b), I briefly explain the ROB procedure applied to cross-sectional time-invariant data potentially affecting labour market outcomes. The starting model to make a move on inference is:

$$Y_i = \sum_{k=1}^{m} \theta_k X_{ik} + \epsilon_i \tag{1}$$

where Y_i is a $N \cdot 1$ vector denoting the variable of interest, with $i = 1, 2, \ldots, N$, $X_{ik} = X_{i1}$, $X_{i2}, \ldots, X_{im}$ is a $[N \cdot m]$ matrix including a few or large set of continuous and/or discrete covariates, with $k = 1, 2, \ldots, m$, $\theta_k = (\theta_1, \theta_2, \ldots, \theta_m)'$ is a $k \cdot 1$ vector of unknown regression coefficients, and $\epsilon_i \sim N(0, \sigma^2)$ is a $N \cdot 1$ vector of disturbances, with σ to be an unknown positive scalar. Here, for simplicity, I drop the constant term and assume that the error component is independent and identically distributed ($i.i.d.$) and homoskedastic.

The main thrust of ROB procedure accounts for providing the $top\ best$[8] model solution (or combination of predictors) better explaining and thus fitting the data. It is very useful when studying the causal link between two or more events affected by additional factors to be involved in the system. In this context, a standard variable selection approach would exclude all the covariates not improving prediction and thus over-confident inferences and decisions about quantities of interest. A practical Bayesian solution to these problems involves estimating 2^m distinct regression models and averaging over them. This is known as BMA and is widely used in the literature to account for model uncertainty (see, e.g., Madigan and Raftery (1994), Madigan et al. (1995), Raftery et al. (1995, 1997)). Nevertheless, the lack of such approach is to use non-informative (or diffuse) priors and 'common' informative priors estimating the unkown regression coefficients θ_k

[8] See, for instance, Pacifico (2020b).

and variance σ^2. The use of Conjugate Informative Proper (CIP) priors[9] in multiple model class – implicit in ROB procedure – overtakes such a limit for three reasons: (i) they change among common parameters entailed in different model solutions; (ii) the distribution of these common parameters change in a corresponding fashion; and (iii) more weight according to model size is assigned. Thus, each possible model solution will be considered likely to be exactly true, without introducing penalty terms or restrictions on data-supported models when there is no relationship between potential predictors. In this way, one will be sure to account for the only relevant factors improving the relationship between obesity and labour market outcomes and thus discard redundant (or non-relevant) variables within the system.

The variable selection problem is addressed by using two auxiliary indicator variables as in Pacifico (2020b). The first corresponds to a vector χ_k, containing every possible 2^m subset choices, with $\chi_k = 0$ if θ_k is small (absence of k-th covariate in the model) and $\chi_k = 1$ if θ_k is sufficiently large (presence of k-th covariate in the model). The second is a vector β_k, corresponding to the regression parameter θ_k when it is sufficiently large (presence of the predictor X_k in the procedure); conversely, the predictor X_k will be ruled out from the procedure.

Let $\mathscr{F}$ be the full model class set[10], the ROB procedure entails shrinking both the model space and the parameter space by matching all potential candidate models in order to jointly deal with overestimation of effect sizes (or individual contributions) and model uncertainty (implicit in the procedure). The shrinking is conducted according to the probability of the candidate models to perform the data, named Posterior Model Probability (PMP) as well. It can be defined as:

$$f(M_k|Y) = \frac{f(M_k)\cdot f(Y|M_k)}{\sum_{M_k\in\mathscr{M}} f(M_k)\cdot f(Y|M_k)} \tag{2}$$

where M_k denotes a countable collection of candidate models containing the vector of the unknown parameters θ and $f(Y|M_k) = \int f(Y|M_k,\theta_k)\cdot f(\theta_k|M_k)d\theta_k$ is the marginal likelihood, with $f(\theta_k|M_k)$ denoting the conditional prior distribution of θ_k. In our context, with both N and T large, the calculation of the integral $f(Y|M_k)$ in not immediate and thus a MC^F integration[11] is involved in the procedure.

After integrating the shrinking, a pool of *best* submodels $M_{\tilde{k}}$ is obtained containing $X_{i\tilde{k}}$ covariates, with $\tilde{k} = 1,2,\ldots,\tilde{m}$, and $M_{\tilde{k}} \ll M_k$ and $\tilde{k} \ll k$ by construction. The first step of TSDPD comes

[9]See, for instance, Pacifico (2020b) for more details on the prior specification strategy.

[10]It contains all the (potential) model solutions. See, for instance, Pacifico (2020b).

[11]More precisely, observations from the joint posterior distribution $f(M_k,\theta_k|Y)$ of (M_k,θ_k) for estimating $f(M_k|Y)$ and $f(\theta_k|M_k,Y)$ are generated recursively. See, for instance, Pacifico (2020b).

to a conclusion once a further shrinkage[12] is conducted in order to obtain a smallest final subset of *top best* submodels (M_ξ) containing the only *significant* solutions contained in the reduced class set $\mathscr{E}$, with $M_\xi \ll M_{\tilde{k}}$. The final regression model will have the form:

$$Y_i = \sum_{\xi=1}^{\kappa} \theta_\xi X_{i\xi} + \eta_i \tag{3}$$

where $X_{i\xi} = X_{i1}, X_{i2}, \ldots, X_{i\kappa}$ is a subset of $X_{i1}, X_{i2}, \ldots, X_{i\tilde{m}}$, with $\xi = 1, 2, \ldots, \kappa$ denoting a subparameter index sufficiently smaller than $\tilde{k}$ ($\xi \ll \tilde{k}$) by construction, θ_ξ denotes the unknown parameters belonging to M_ξ, which contains the only *significant* solutions, and η_i is the *i.i.d.* error term.

Finally, the exact and final solution will correspond to one of the submodels M_ξ with higher log natural Bayes Factor (lBF):

$$lBF_{\xi,\tilde{k}} = log\left\{\frac{\pi(M_\xi | Y_n = y_n)}{\pi(M_{\tilde{k}} | Y_n = y_n)}\right\} \tag{4}$$

In this empirical analysis, the lBF will be interpreted through the scale of evidence according to a generalised version of Kass and Raftery (1995), as in Pacifico (2020b):

$$\begin{cases} 0 < lB_{\xi,\tilde{k}} \leq 2 & \text{no evidence for submodel } M_\xi \\ 2 < lB_{\xi,\tilde{k}} \leq 6 & \text{moderate evidence for submodel } M_\xi \\ 6 < lB_{\xi,\tilde{k}} \leq 10 & \text{strong evidence for submodel } M_\xi \\ lB_{\xi,\tilde{k}} > 10 & \text{very strong evidence for submodel } M_\xi \end{cases} \tag{5}$$

2.2 Dynamic Panel Data with GMM Estimators - Second Step

The baseline TSDPD model is:

$$Y_{it} = \delta_i^{'} + \sum_{r=1}^{\rho} \gamma_r^{'} W_{it-r} + \sum_{l=0}^{\lambda} \sum_{\xi=1}^{\kappa} \theta_{l\xi}^{'} X_{it-l,\xi} + u_{it} \tag{6}$$

where Y_{it} is a $NT \cdot 1$ vector of outcomes, δ_i is a $N \cdot 1$ heterogeneous intercept, W_{it-r} is a $NT \cdot 1$ vector of predetermined variables, $X_{it-l,\xi}$ is a $NT \cdot \kappa$ matrix containing continuous/discrete endoge-

[12] It refers to the second stage concerning the ROB procedure. See, for instance, Pacifico (2020b).

nous variables, with $l = 0,1,2,\ldots,\lambda$, $r = 1,2,\ldots,\rho$ denotes generic Auto-Regressive (AR) orders for the predetermined variables, γ_r and $\theta_{\tilde{\lambda}\xi}$ are the autoregressive coefficients to be estimated for each i and couple of (i,ξ), with $\tilde{\lambda} = 1,\ldots,\lambda$, and $u_{it} \sim i.i.d.N(0,\sigma_u^2)$ is a $NT \cdot 1$ vector of unpredictable shock (or idiosyncratic error term), with $E(u_{it}) = 0$ and $E(u_{it} \cdot u_{js}) = \sigma_u^2$ if $i = j$ and $t = s$, and $E(u_{it} \cdot u_{js}) = 0$ otherwise.

Here, some considerations are in order: (i) the predetermined variables contain the lagged values of the outcomes Y_{it} and lags of heterogeneous individual-specific factors; (ii) the δ_i's denote cross-unit heterogeneity affecting the outcomes Y_{it}; (iii) a correlated random effects approach is adopted in which the δ_i's are treated as random variables and possibly correlated with some of the covariates within the system; (iv) the roots of $r(B) = 0$ and $\tilde{l}(B) = 0$ lie outside the unit circle so that the AR processes implicit in the model (6) are stationarities, with $\tilde{l} = 1,2,\ldots,\tilde{\lambda}$ denoting generic AR orders for the endogenous variables and B referring to the lag operator; and (v) the instruments are fitted values from autoregressive parameters based on all available lags of time-varying variables and their causal interactions.

A common model building strategy is to select the exact differentiation order and thus plausible values of AR lag orders on statistics calculated from the data to assess the stationarity of the processes implicit in (6) for each sample unit. In this study, I use the Schwarz Bayesian Information Criterion (SBIC) – displayed in equation (7) – to select the optimal lag length in AR time-series and the Augmented Dickey-Fuller (ADF) test – displayed in equation (8) and stacked for i – to choose the order of integration to ensure stationarity.

$$BIC(\dot{p}) = log(\hat{\sigma}_u^2) + \frac{(\dot{p}) \cdot log(T)}{T} \tag{7}$$

$$\Delta Y_t = \mu + \varrho t + \vartheta Y_{t-1} + \varphi_1 \Delta Y_{t-1} + \ldots + \varphi_{\dot{p}-1} \Delta Y_{t-\dot{p}+1} + \varepsilon_t \tag{8}$$

where $\hat{\sigma}_u^2$ denotes the Maximum Likelihood Estimate (MLE) of σ_u^2, μ is a constant, ϱ is the coefficient on a time trend, and $\dot{p} = (\rho, \tilde{\lambda})$ denotes the lag orders of the AR processes in model (6).

Let the stationarity hold in the system, the time-series regressions are valid (or computational) and GMM estimators are feasible. Here, the choice of lag periods is critical, because too few lags provoke autocorrelated errors and thus spurious test statistics, while too many lags reduce the power of the test. In this context, the choice of lag periods obeys the rule of Dumitrescu and

Hurlin (2012), which says that the minimum time extent for ρ and $\tilde{\lambda}$ should be chosen according to $T > 5+2p$ (where T is the number of time periods and p is the number of general lags). Then, since the computation of GMM estimators requires restrictions on the initial conditions process, I assume that δ_i and $u_{i,t}$ are independently distributed across i and have the familiar error components structure:

$$E(\delta_i) = 0, E(u_{it}) = 0, E(u_{it} \cdot \delta_i) = 0 \qquad for \quad i = 1, \ldots, N \quad and \quad t = 2, \ldots T \tag{9}$$

and

$$E(u_{it} \cdot u_{is}) = 0 \qquad for \quad i = 1, \ldots, N \quad and \quad t \neq s \tag{10}$$

In addition, I also assume the standard assumption concerning the initial conditions $Y_{i,t=1}$ (see, e.g., Ahn and Schmidt (1995)):

$$E(Y_{i,t=1} \cdot u_{it}) = 0 \qquad for \quad i = 1, \ldots, N \quad and \quad t = 2, \ldots T \tag{11}$$

Conditions (9), (10), and (11) imply moment restrictions that are sufficient to address exact identification in a context of random effects and estimate γ_r and $\theta_{\tilde{l}}$ for $T \geq 3$.

Accounting for time-varying and endogenous variables, I build on Dumitrescu and Hurlin (2012) again to test for the existence of Granger causality in heterogeneous dynamic panels between the system's covariates and the outcomes, and vice versa. Under the null hypothesis, there is no causal relationship for any of the units of the panel (Homogeneous Non Causality hypothesis), whereas there is a causal relationship from $X_{it-l,\xi}$ to $Y_{i,t}$ for a subgroup of units (Heterogeneous Non Causality hypothesis) under the alternative. In a time-series context, the standard causality tests consist in testing linear restrictions on the slope parameters in model (6). One must be very careful to the issue of heterogeneity of the parameters since it directly affects the paradigm of the representative agent and thus the conclusions with respect to causality relationships. It is well known that the estimates of autoregressive parameters obtained under the wrong hypothesis are biased (see e.g., Pesaran and Smith (1995)). Then, if one imposes the homogeneity of coefficients, the causality test-statistics can lead to fallacious inference. Intuitively, the estimators obtained in a homogeneous model will converge to a value close to the average of the true coefficients, and if this mean is itself close to zero, one risks to accept at wrong the hypothesis of no causality. In this analysis, the optimal lag length to test Granger-causality has been set by using the Arellano's test[13].

Stacking for i, three main findings are in order. First, the GMM estimators ($\hat{\gamma}_r$ and $\hat{\theta}_{lk}$) will be consistent and unbiased accounting – by construction – for endogeneity issues, structural model uncertainty, and (Granger-)causality in dynamic panels. Second, they will also be able to investigate – by assumption – the presence of relevant interconnections and interdependencies between Y_t and X_t, and between X_t and its lags[14]. Third, all the variables within the system will be – by construction – potentially $significant$ with highly strong predictive accuracy.

3 The Empirical Application: Evidence across Italian Regions

3.1 Literature Review and Discussion with Related Works

Obesity is a complex condition that has serious health, social, and psychological dimensions, affecting all ages and socioeconomic groups. The negative impacts of obesity on health are well known: obesity is a major contributor to the global burden of chronic disease and disability, including diabetes, cardiovascular disease, and cancer.

The impact of excess weight in the workplace has also been a domain of investigation, with a lot of studies highlighting the increasing prevalence of obesity across industries and occupational groups negatively affects employment and wages (see, e.g., Morris (2007), Tunceli et al. (2006), Mosca (2013), Caliendo and Lee (2013), and Lundborg et al. (2010)). Although preliminary studies suggest that obesity may differentially affect work productivity and costs, based on occupational requirements, there is also substantial evidence that obese people, particularly women, are less likely to be employed and, when employed, are likely to earn lower wages due to employer discimination (see, e.g., Averett and Korenman (1996), Harper (2000), Loh (1993), and Pagan and Davila (1997)).

The literature on the possible links between obesity and adverse labour market outcomes has been growing since the mid-1990s. The increased consumption of more energy-dense foods and foods with high levels of sugar and saturated fats, combined with reduced physical activity, have

[13] See, for instance, Arellano (2003).

[14] Similar frameworks, with appropriate Bayesian empirical specifications, have been used to make inference and obtain posterior distributions among time-varying macroeconomic-financial variables in multicountry panel data (see, e.g., Canova and Ciccarelli (2009), Canova et al. (2007), and Pacifico (2019b,c)).

led to obesity rates that have risen significantly since 1980 in developed (USA, UK, Australia), transition (Eastern Europe), and emerging (the Middle East, China) economies. From a policy perspective, prevention and effective public policies need to be accounted for understanding whether obesity is associated with adverse labour market outcomes and establishing the risk factors associated with these outcomes. From a modeling perspective, there is an active debate about whether the relationship between labour market outcomes and obesity are or not due to causal link. For example, people who are paid less might become obese in part because they cannot afford healthful food and must rely instead on low-cost, low-nutrition, calorie-dense foods (see, for instance, Barnay (2015) and Gortmaker et al. (1993)). Some evidence shows that non-employment and poor working conditions have detrimental effects on health and a lack of control over the amount of time devoted to work (see, for instance, Datta and Nicolai (2008), Barnay (2015), and Llena-Nozal (2009)), and effects of problem drinking on employment (see, for instance, Mullahy and Sindelar (1993), Stuckler et al. (2009), and Marchand et al. (2011)).

The relationship between high body weight and labour market outcomes has been primarily studied by using data from developed and high income countries, such as the US and West Europe (e.g., England, Denmark, and Finland). The main labour market outcomes studied were wages/earnings, employment, and occupational selection (see, e.g., Chou et al. (2004), Lakdawalla et al. (2005), Rashad et al. (2006), Burkhauser and Cawley (2008), Burkhauser et al. (2009), Komlos and Brabec (2010), Flegal et al. (1998), and Flegal et al. (2010)). Earlier papers focused on the US have used the National Longitudinal Survey of Youth (NLSY) data, and found mixed results (see, e.g., Register and Williams (1990), Loh (1993) and Pagan and Davila (1997)). Some shortcomings of these studies are that they ignore the potential endogeneity of obesity – making causal inference impossible – account for small and unrepresentative samples, and estimate cross-sectional data.

Later studies have tried to address endogeneity issues because of hidden[15] or hard-to-measure factors that might affect both obesity, defined as a Body Mass Index (BMI)[16], and labour market outcomes. For example, Cawley and Chad (2012) use an Instrumental Variable (IV) method to estimate the impact of obesity on medical costs in order to deal with endogeneity problems and thus reduce the empirical bias of estimates. The main thrust of the IV model has been to put more emphasis on the causal effect of obesity on medical care costs in contrast to previous studies

[15] Hidden factors are variables that are not directly observed but are rather deduced from other variables that are observed and thus directly measured.

[16] Body Mass Index is defined as weight in kilograms divided by height in meters squared.

focusing on their correlation and thus overestimating the causal relationship. Nevertheless, the only inclusion of $internal$ instruments[17] makes the analysis unable to investigate additional factors affecting the (causal) link between weight and labour market outcomes because of chronic diseases. Another related important work has been developed by Baum and Ford (2004). They use NLSY data to investigate the effects of obesity on wages by gender accounting for a (potential) set of individual characteristics. The results are consistent with the recent literature which recognizes that an obesity wage penalty persists for both males and females, even if this penalty seems to be larger for females. Essentially, in this context, individuals serve as their own control in fixed effects models. They use two different sets of individual background characteristics: time-invariant individual-specific heterogeneity and time-varying family-specific heterogeneity. However, if these unobservable factors vary over time and/or differ across units, individual fixed-effects models cannot account for them and thus the corresponding estimates will be biased as well. In addition, possible correlations between the lags in the deterministic and causal components need to be accounted for.

3.2 Data Description and Preliminary Analysis

The data are collected referring to two databases: (i) the Central Institute of Statistics (ISTAT) and (ii) the report on equitable and sustainable well-being (BES). The former is an Italian public research body dealing with general population censuses, services and industry, agriculture, household sample surveys, and general economic surveys at national level. The BES is not just an editorial product, but a line of research and thus a process that takes the multidimensionality of well-being as a starting point and describes – in a comprehensive way – the quality of life in Italy. Every project to measure equitable and sustainable well-being aims at evaluating the progress of society from either an economic or a social and environmental point of view.

In this study, I account for all the 21 Italian regions and use a time-series spanning the period 2007 – 2017, accounting for a large pool of (potential) indicators such as gender, education level, high body weight, wage, and other factors related to labour market dynamics and individual characteristics.

By running the first step of the TSDPD estimation (Section 2.1), the dataset collects 14 (poten-

[17] They use the weight of a biological relative as instrument for the respondent's weight. It has been used in previous studies to assess the impact of weight on other outcomes such as wages (see, for instance, Cawley (2004) and Kline and Tobias (2008)).

tial) candidate predictors affecting wage effects which denotes the initial[18] variable of interest (Table 1). All data have been taken in percentage or logarithm with respect to their measurement unit, and there are $2^{14} = 16,384$ possible model solutions (M_k) in the system (1). In this context, two further statistics are accounted for. The Posterior Inclusion Probabilities (PIPs), corresponding to the sum of the PMPs displayed in equation (2), for all M_k models wherein a covariate X_k has been included with the auxiliary variable $\chi = 1$, and the Conditional Posterior Sign (CPS) for the sign certainty, taking values close to 1 or 0 if a covariate X_k has a positive or negative effect on wage, respectively.

Here, some considerations are in order: (i) predictors (7, 8, 9, 10, 12, 13) denote heterogeneous individual characteristics possibly affecting the covariates within the system; (ii) predictors (2, 3, 4, 5, 11) denote the risk and socioeconomic factors; (iii) predictors (1, 6, 14) denote other (potential) endogenous variables; (iv) the added value per employee[19] is used to estimate cross-unit wage effects among Italian regions; and (v) the predictor 2 stands for excess weight (obesity) in terms of BMI measured as weight/height[20]. All the variables within the system are time-varying.

Without accounting for GMM estimates, preliminary findings are addressed. Some individual characteristics and socioeconomic factors such as predictors (3, 4, 5, 12, 13) – oftentimes overlooked on the literature – show a very strong impact on the initial variable of interest ($wage$). More precisely, family relationship and free time satisfaction tend to positively affect wages in contrast with smoking[21] and alcohol consumption, and sedentary rate. Experience (predictors 7 and 8) and employment insurance (predictors 6 and 10) tend to be positively associated with wages in contrast with inactivity (predictor 9) and risk of poverty (predictor 11) rate. The impact of excess weight (predictor 2) and weighted income (predictor 1) tend to have highly large wage penalties[22] and opportunities with a CPS close to 0 and 1, respectively. Finally, the gender indicator (predictor 14) need to be deepened showing an ambiguous sign certainty.

By looking into which covariates are included with higher frequency in the submodels solutions ($M_{\tilde{k}}$), the first shrinking is conducted. In this analysis, 11 $best$ covariates are found, obtaining $2^{10} = 1,024$ $best$ model solutions (Table 2).

[18]Other potential variables affecting labour market outcomes will be addressed later by conducting the Bayesian framework in Section 2.1.

[19]Source: ISTAT database.

[20]Source: ISTAT database.

[21]See, for instance, Levine et al. (1997).

[22]See, for instance, Lundborg et al. (2010).

Table 1: Dataset

Idx.	Predictor	Label	Unit	PIP (%)	CPS
1	weighted income per capita	($income$)	thousands € (log.)	33.04	1.00
2	overweight (obesity)	(obe)	std. rates per 100 people	82.24	0.00
3	consumption of tobacco	($smoke$)	std. rates per 100 people	27.05	0.21
4	consumption of alcohol	($alcohol$)	std. rates per 100 people	33.68	0.30
5	sedentary rate	(sed)	std. rates per 100 people	8.25	0.31
6	employment rate	($employ$)	% values	25.61	0.92
7	high school diploma	($school$)	% values	0.96	0.87
8	graduates/other qualifications	($degree$)	% values	9.85	0.89
9	neither studing nor working	(nsw)	% values	0.83	0.29
10	fixed-term contract (= 5 years)	($fterm$)	% values	0.40	0.63
11	risk of poverty	(rop)	% values	2.79	0.09
12	family relationship satisfaction	($family$)	% values	10.50	0.63
13	free time satisfaction	($ftime$)	% values	8.86	0.65
14	indicator variable for gender	($gender$)	$[0,1]$	7.15	0.48
-	added value per employee	($wage$)	thousands € (log.)	-	-

The Table is so split: the first column denotes the predictor number; the second and the third column describe the predictors and the corresponding labels; the fourth column refers to the measurement unit; and the last two columns displays the PIPs (in %) for each predictor and the CPS, respectively. The last row refers to the initial variable of interest. The contraction $std.$ stands for 'standardized'. All data refer to ISTAT and BES databases.

Table 2: $Best$ Potential Combination of Predictors for Wages

Predictor	Idx.	PIP (%)	CPS
$income$	1	14.16	1.00
obe	2	46.84	0.00
$smoke$	3	13.03	0.40
$alcohol$	4	25.56	0.37
sed	5	0.11	0.57
$degree$	8	22.26	1.00
rop	11	0.24	0.48
$family$	12	23.12	0.78
$ftime$	13	1.24	0.82
$gender$	14	7.52	0.54

The Table is so split: the first two columns refer to the predictors and the corresponding index number defined in Table 1, and the last two columns display the PIPs (in %) and the CPS, respectively.

Entailing a further shrinking as involved in ROB procedure, the predictors (1, 2, 3, 4, 8, 12, 13, 14) would look like the $top\ best$ combination of covariates (X_ξ) with higher PIPs[23]. Here, three main findings are addressed. First, the model uncertainty and overfitting implicit in the ROB procedure are avoided: indeed, the sign certainty tends to be close to 0 – such as for predictor

[23] More precisely, the $top\ best$ covariates are selected with a PIP ≥ 0.5% for a sufficient prediction accuracy in explaining the data. See, for instance, Pacifico (2020b).

2 – and 1 – such as for predictors 1, 8, 12 and 13. Uncertain effects persist in predictors 3, 4, and 14. Thus, they should be interpreted with caution. For example, accounting for alcohol and smoking consumption, heavy smokers and drinkers – most likely – would be negatively associated with wages. In the matter of gender indicator, males would likely be associated with lower wage penalties. However, this predictor need to be assessed in depth in order to highlight different dynamics between males and females.

Finally, according to the log Bayes Factor in equation (4), the final solution[24] better performing the data corresponds to the model consisting of predictors $(1, 2, 3, 4, 8, 12, 13, 14)$. More precisely, the time-invariant version of the model (6) is:

$$\begin{aligned} wage_i &= c + \theta_1 obe_{i1} + \theta_2 income_{i2} + \theta_3 smoke_{i3} + \theta_4 alcohol_{i4} + \theta_5 degree_{i5} + \\ &\quad + \theta_6 family_{i6} + \theta_7 ftime_{i7} + \theta_8 gender_{i8} + \tilde{\eta}_i \end{aligned} \tag{12}$$

where c is an intercept, θ_ξ denotes the unknown parameters belonging to M_ξ, with $\xi = 1, 2, \ldots, 8$, and $\tilde{\eta}_i$ is a $N \cdot 1$ vector containing the $i.i.d.$ disturbances, with $\sigma_{\tilde{\eta}_i}$ to be an unknown positive scalar.

Before moving forward with the second step which involves the dynamic version of model (12), endogeneity issues between wages and some covariates $X_{\tilde{k}}$ dropped in the first shrinking (Table 2) need to be clarified. For example, labour participation rate (predictor 6) – with highly larger PIP than the other discarded predictors – would be an interesting instrument to investigate how employment prospects affect causal link between labour market outcomes and obesity. According to Pacifico (2020b), a Two-Stage Least Squares (TSLS) estimator is used to solve such endogeneity problems when Z instruments occur, with Z denoting a $[N \cdot \dot{m}]$ matrix of instruments. In this context, the validity of instrument is addressed by constructing a Bayesian test of the identification restrictions based on model averaged posterior predictive p-values.

The auxiliary regression of model (12) is:

[24] More precisely, looking into which models included perform better the data, 22 $top\ best$ model solutions (M_ξ) have been found. The higher lBF – associated with the final solution – equals 9.17.

$$income_i = \tilde{c} + \pi_1 employ_{i1} + \pi_2 obe_{i2} + \pi_3 smoke_{i3} + \pi_4 alcohol_{i4} + \pi_5 degree_{i5} + $$
$$+ \theta_6 family_{i6} + \theta_7 ftime_{i7} + \theta_8 gender_{i8} + \nu_i \quad (13)$$

where $\tilde{c}$ is an intercept, $\pi_.$ denotes the unknown parameters of (13) and ν_i is the $N \cdot 1$ vector of disturbances independent and identically distributed with respect to $\tilde{\eta}_i$, with σ_{ν_i} to be an unknown positive scalar.

The Table 3 summarizes the results. Here, some considerations are addressed: (i) only one variable (predictor 6) serves as instrument (Z_1); (ii) the negative impact of obesity (predictor 2) and smoking and alcohol consumption (predictors 3 and 4) tends to increase when the instrument Z_1 is accounted for; (iii) positive effects associated with family relationship (predictor 12) and free time satisfaction (predictor 13) tend to be unvaried; (iv) levels of experience (predictor 8) tends to show a higher positive explanatory power when the instrument Z_1 occurs; and (v) there is a difference between men's and women's earnings. These findings highlight that employment rate (predictor 6) would work as potential instrument investigating and clarifying the causal link between obesity and labour market outcomes.

Finally, the posterior predictive p-values for both equations (12) and (13) are close to zero and thus model assumptions are appropriately identified (Table 3). All of them will be discussed in depth and jointly verified in the second step of the TSDPD estimate, which corresponds to the main thrust of this study.

Table 3: Causality - A First Investigation

Predictor	Model (12)	Model (13)	Effect
obe	(0.042)**	(0.006)***	(-)
income	(0.054)*	-	(+)
smoke	(0.059)*	(0.008)***	(-)
alcohol	(0.026)**	(0.031)**	(-)
degree	(0.037)**	(0.007)***	(+)
family	(0.047)**	(0.031)**	(+)
ftime	(0.035)**	(0.041)**	(+)
gender	(0.026)**	(0.018)**	(+)
employ	-	(0.002)***	(+)
	$\bar{R}^2_{M12} = 65.43$	$\bar{R}^2_{M13} = 63.36$	
	$\varpi_{\xi,M12} = 0.00$	$\varpi_{\xi,M13} = 0.00$	

The Table is so split: the first column refers to predictors; the second and third column display the estimates in terms of p-values with the corresponding significant codes; and the fourth column displays the wage effects. The last two rows refer to $\bar{R}^2$ and posterior predictive p-values (ϖ) for both the models. The significant codes are: *** significance at 1%, ** significance at 5%, and * significance at 10%.

3.3 Empirical Results and Prevention Policy-Relevant Strategies

Let on the *top best* combination of predictors in model (12), the dynamic model (6) can be assessed. Before estimating it, one needs to choose the optimal lag of the time-series and ensure their stationarity in order to be sure that their distribution neither follows any trend nor changes over time. The latter is a key requirement for the validity of time-series regressions. The Augmented Dickey-Fuller method in (8) is used with the null hypothesis that all panels contain a unit root, saying that the series are non-stationary. The alternative hypothesis of stationarity is accepted if the probability is less than the critical value 0.05.

If T is small (e.g., $T < 10$), it is a manageable number and thus restrictions on the number of past lags used are not necessary. Conversely, if T is fairly large (e.g., $T \geq 10$, just as in our case), an unrestricted set of lags will introduce a huge number of instruments, with a possible loss of efficiency. By using the lag limits options, one may specify – for example – that only lags $[2-5]$ are to be used in constructing the GMM instruments. Given such solution to the tradeoff between lag length and sample length, the Holtz-Eakin et al. (1988)'s suggestion can be followed by including all available lags of the untransformed variables as instruments. For endogenous variables, lags 2 and higher are available. For predetermined variables – that are not strictly exogenous – lag 1 is also valid since its value is only correlated with errors dated $t-2$ or earlier.

By equations (7) and (8), I estimate 8 different AR processes to obtain potential instruments as 'GMM-style'. They are constructed by following the Arellano and Bond (1991)'s logic, making use of multiple lags. In Table 4, I display the AR time-series, the ADF tests in terms of p-values, and Ljung-Box test statistics of the series to jointly assess the robustness of the estimates and investigate linear dependencies among series. All the series are stationary and valid, showing highly strong linear dependencies and no autocorrelation among residuals. Thus, unobserved heterogeneity and model misspecification problems matter. Here, the maximum differencing order to test stationarity sets 1 for all the predictors within the system.

Table 4: AR Processes and Diagnostic Tests

Predictors	`wage`	`obe`	`employ`	`smoke`	`alcohol`	`degree`	`family`	`ftime`
$AR(p)$	(3)	(3)	(3)	(3)	(3)	(3)	(3)	(3)
ADF	0.01**	0.01**	0.02**	0.02**	0.03**	0.01**	0.02**	0.01**
LGB_s	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***
LGB_r	0.73	0.90	0.99	0.86	0.92	0.55	0.85	0.90

The Table is so split: the first row refers to the *top best* predictors and the lagged outcomes; the second row accounts for AR(p) models, with p denoting the optimal lag; the third row stands for the ADF tests in terms of p-values; and the last two rows stand for Ljung-Box test statistics of the series (LGB_s) and residuals (LGB_r) in terms of p-values. The significant codes are: *** significance at 1%, ** significance at 5%, and * significance at 10%.

The TSDPD model in (6) can be written as:

$$Y_{it} = \delta_i^{'} + \sum_{r=1}^{3} \gamma_r^{'} W_{it-r} + \sum_{\tilde{l}=1}^{3} \sum_{\xi=1}^{4} \theta_{\tilde{l}\xi}^{'} X_{it-\tilde{l},\xi} + \sum_{\xi=1}^{4} \theta_{\xi}^{'} X_{it,\xi} + u_{it} \tag{14}$$

where δ_i is a $N \cdot 1$ heterogeneous intercept observed in t, W_{it-r} is a $NT \cdot 1$ vector of predetermined variables containing lagged labour market outcomes[25] and lagged heterogeneous individual characteristics[26], with $\rho = 3$, $X_{it,\xi}$ and $X_{it-\tilde{l},\xi}$ are $NT \cdot \kappa$ matrices containing continous and discrete endogenous variables[27] in t and their corresponding lagged values in $t - \tilde{l}$, respectively, with $\kappa = 4$ and $\tilde{\lambda} = 3$, and u_{it} $i.i.d.N(0, \sigma_u^2)$ is a $NT \cdot 1$ vector of idiosyncratic error term.

According to Dumitrescu and Hurlin (2012) and the preliminary results obtained in Section 3.2, in a context of time-series analysis, I investigate the existence of Granger causality in the heterogeneous (balanced) dynamic panel model between: (i) the system's time-varying explanatory variables ($\tilde{W}_{it-r}^{1}$, $X_{it-l,\xi}$) and wage (Y_{it}^{W}), with $\tilde{W}^1$ denoting the lagged individual heterogeneity

[25] They correspond to wage and predictor 6.
[26] They correspond to predictors (8, 12, 13).
[27] They correspond to predictors (2, 3, 4, 14).

and the only lagged employment rate; (ii) wage (Y_{it}^{W}) and the time-varying explanatory variables ($\tilde{W}^{1}_{it-r}$, $X_{it-l,\xi}$); (iii) employment rate (predictor 6) and the time-varying explanatory variables ($\tilde{W}^{2}_{it-r}$, $X_{it-l,\xi}$), with $\tilde{W}^{2}$ denoting the lagged individual heterogeneity and the only lagged wage; and (iv) the time-varying explanatory variables ($\tilde{W}^{2}_{it-r}$, $X_{it-l,\xi}$) and employment rate (predictor 6). Table 5 displays the full results of Granger-causality. The optimal lag length to test Granger-causality has been chosen equal to 3, since it enables to eliminate serial correlation in residuals u_{it}, and the subgroup to be tested corresponds to the model solutions in M_{ξ}.

In summary, I find eleven main distinct Granger-causality relationships: (i) a two-way causal link between excess weight and work; (ii) a two-way causal link between excess weight and wage; (iii) a two-way causal link between work and wage; (iv) a two-way causal link between wage and work; (v) a two-way causal link between smoke and workplace tasks; (vi) a two-way causal link between alcohol use and workplace tasks; (vii) a unique causal link between education and work; $(viii)$ a unique causal link between education and wage; (ix) a unique causal link between family relationship satisfaction and wage; (x) a unique causal link between free time satisfaction and work; and (xi) a two-way causal link between free time satisfaction and wage. The results at points (i), (ii), and (iv) find confirmation with the existing literature (see, for instance, Morris (2007), Tunceli et al. (2006), Lundborg et al. (2010), and Jusot et al. (2008)). Opposing findings hold about causal linkages between smoking and wages (see, for instance, Levine et al. (1997)) and non-causal linkages between alcohol use and workplace tasks (see, for instance, Jarl and Gerdtham (2012)). The results at points (vii) – (xi), dealing with individual-specific heterogeneity, represent one of the main aims of this study. However, by having a look at statistical significance and p-values, I are not able to focus on effect magnitude. Thus, the TSDPD estimation needs to be accounted for. In this context, gender differences and similarities according to predictor 14 – one of the $top\ best$ covariates found in Section 3.2 – are also investigated.

The TSDPD estimates confirm and deepen such findings (Table 6). I split them in two parts – modeling and policy perspective – by running three different models: (i) Model 1 accounting for all sample units; (ii) Model 2 accounting for male individuals; and (iii) Model 3 accounting for female individuals.

From a modeling perspective, according to Model 1, negative (causal) impacts of excess weight run in the workplace (lower probability to be employed) and wage effects (larger wage penalties). The same occurs in the opposite direction: negative health impacts might hold due to excessive amount of time devoted to work or unsatisfactory pay. Socioeconomic factors are negatively re-

Table 5: Granger-Causality

From TEV^1 to Y^W	Obesity	Employ	Smoke	Alcohol	Degree	Family	Ftime
Z-tilde's Test Statistics	6.48*** (0.00)	3.67** (0.04)	0.55 (0.58)	0.87 (0.38)	2.71* (0.08)	0.32 (0.90)	2.57** (0.02)
From Y^W to TEV^1	**Obesity**	**Employ**	**Smoke**	**Alcohol**	**Degree**	**Family**	**Ftime**
Z-tilde's Test Statistics	5.13** (0.03)	3.10** (0.02)	0.82 (0.41)	0.82 (0.32)	0.77 (0.21)	1.99** (0.04)	2.88*** (0.00)
From TEV^2 to $P6$	**Obesity**	**Wage**	**Smoke**	**Alcohol**	**Degree**	**Family**	**Ftime**
Z-tilde's Test Statistics	4.14*** (0.00)	2.75** (0.03)	2.26** (0.02)	3.58*** (0.00)	1.77* (0.08)	0.93 (0.64)	0.72 (0.29)
From $P6$ to TEV^2	**Obesity**	**Wage**	**Smoke**	**Alcohol**	**Degree**	**Family**	**Ftime**
Z-tilde's Test Statistics	3.97*** (0.00)	2.67** (0.03)	2.16** (0.03)	4.45*** (0.00)	0.80 (0.35)	0.58 (0.57)	4.41*** (0.00)

The Table displays all the Z-tilde test statistics and p-values (in parenthesis) on the Granger-causality in the dynamic panel model (14). Here, TEV^1 stands for time-varying explanatory variables ($\tilde{W}^1_{it-r}$, $X_{it-l,\xi}$) including employment rate, TEV^2 stands for time-varying explanatory variables ($\tilde{W}^2_{it-r}$, $X_{it-l,\xi}$) including wage, Y^W refers to wage, and $P6$ refers to employment rate. The significant codes are: *** significance at 1%, ** significance at 5%, and * significance at 10%.

lated to wage effects and working conditions. More precisely, heavy smokers and drinkers tend to show negative (causal) effects on employment (likely to be unemployed) and negative (non-causal) impacts on earnings (lower wages). Thus, the former would be associated with the risk of recurrent sickness leave causing long-term absence from work, increasing welfare payments for the treatment of these diseases, and increasing probability of early retirement from the labour force/unemployment (causality between work and wage and vice versa). Contrary, wage effects related to smoking and alcohol use would depend on other not-directly observed factors such as ability, working hours, and employment status. Wage improvements strongly depend on family relationship satisfaction and even more on free time activities (two-way causal link); whereas, positive (causal) impacts of good work performance would only depend on lifestyle. Earnings and working conditions are highly and positively correlated between them and – at the same time – affected from socioeconomic circumstances and individual-specific heterogeneity (e.g., tobacco smoking, alcohol use, family relationship and free time satisfaction, and employment prospects). Finally, positive (causal) impacts of education level run in both the workplace (better employment prospects) and wage effects (high possible wage improvements).

Accounting for gender (Models 2 and 3), three main results are in oder. First, obesity wage and employment penalties persist for males and even more for females, just as in the previous studies (see, e.g., Cawley and Chad (2012), Baum and Ford (2004), and Flegal et al. (2010)). Second, socioeconomic factors negatively affect employment opportunities and wage effects for male and slightly more for female individuals. Here, a significant difference is found for smoking and alcohol use, where negative impacts in working conditions and wage effects matter more for females. Third, education level, employment opportunities and wage improvements seem to matter more among male individuals.

In summary, according to the overall estimates in Table 6, all the variables (predetermined and endogenous factors) and individual-specific heterogeneity are significant and the time-series results are robust and valid (no autocorrelation among residuals and highly strong linear dependencies). These findings highlight three main conclusions: (i) the usefulness to address variable selection problems by dealing with endogenity issues and structural model uncertainty; (ii) the performance of the TSDPD model in improving causal relationships between obesity and labour market outcomes; and (iii) the importance to account for heterogeneous individual characteristics and socioeconomic factors when studying causal links in dynamic panel setups.

From a policy perspective, appropriate prevention and health care in support of chronic diseases might lead to consistent gains in economic production through healthier workplace and more active workforce. In this context, prevention policy-relevant strategies designed to deal with key behavioural risk factors such as obesity, smoking and alcohol consumption, and negative socioeconomic factors would be able to increase employment opportunities and wage effects, improve labour productivity, and reduce social disparities in health and – possibly – in gender. Moreover, in contrast with the existing literature[28], this study highlights that adverse labour market outcomes and thus associated production losses depict high costs society in terms of additional cost components. The highest component is related to individuals with highly negative socioeconomic factors and large risk of chronic diseases in terms of labour force. Most of fiscal revenues addressed to public expenditures will be assigned to increase employment opportunities, improve working conditions, and employ welfare spending. It follows that employers will support temporary replacement costs and recurring staff turnover by implying competitive losses in the labour market. The same occurs with other socioeconomic factors such as smoking and alcohol consumption, which are associated with adverse labour market outcomes by involving additional costs for employers as well as workers.

[28] Existing studies assume that production losses are associated with adverse labour market outcomes (employment and wage penalties) and current labour costs reflect long-term absence from work (e.g., due to early exit from the labour force, early retirement, and unemployment).

Table 6: TSDPD Estimation - Labour Market Outcomes

Variables	Model 1 - Total	Model 2 - Male	Model 3 - Female
	Wage Effects		
L^j.wage	$0.43***$ (0.07)	$0.51***$ (0.07)	$0.46***$ (0.07)
obesity	$-0.22***$ (0.20)	$-0.11***$ (0.16)	$-0.28***$ (0.17)
employ	$0.64***$ (0.18)	$0.70***$ (0.15)	$0.42***$ (0.16)
smoke	$-0.44*$ (0.24)	$-0.24*$ (0.17)	$-0.51*$ (0.21)
alcohol	$-0.47*$ (0.17)	$-0.17*$ (0.11)	$-0.20*$ (0.20)
degree	$0.37***$ (0.15)	$0.28***$ (0.14)	$0.23***$ (0.10)
family	$0.14**$ (0.13)	$0.13**$ (0.12)	$0.16**$ (0.12)
ftime	$0.24***$ (0.15)	$0.30***$ (0.13)	$0.28***$ (0.14)
Q_S	0.00***	0.00***	0.00***
Q_A	0.01**	0.01**	0.00***
Q_{LB}	0.00***	0.00***	0.00***
N	171	171	171
	Employment Rate		
L^j.employ	$0.63***$ (0.06)	$0.72***$ (0.06)	$0.54***$ (0.07)
obesity	$-0.10***$ (0.07)	$-0.24***$ (0.06)	$-0.31***$ (0.06)
wage	$0.45***$ (0.02)	$0.48***$ (0.02)	$0.32***$ (0.03)
smoke	$-0.28**$ (0.08)	$-0.11*$ (0.06)	$-0.22**$ (0.09)
alcohol	$-0.22**$ (0.05)	$-0.15**$ (0.04)	$-0.18**$ (0.08)
degree	$0.25***$ (0.04)	$0.18***$ (0.05)	$0.14***$ (0.03)
family	$0.16**$ (0.04)	$0.14***$ (0.04)	$0.17***$ (0.06)
ftime	$0.10*$ (0.05)	$0.09***$ (0.05)	$0.11**$ (0.05)
Q_S	0.00***	0.00***	0.00***
Q_A	0.00***	0.00***	0.02**
Q_{LB}	0.00***	0.00***	0.00***
N	171	171	171

Here, the Standard Errors, in parenthesis, are adjusted for the heteroskedasticity and L^j stands for the lag operator, with $j = 3$. The instruments used to estimate the TSDPD in equation (14) are: $obesity_{t-1,t-2,t-3}$, $employ_{t-1,t-2,t-3}$, $wage_{t-1,t-2,t-3}$, $smoke_{t-1,t-2,t-3}$, $alcohol_{t-1,t-2,t-3}$, $degree_{t-1,t-2,t-3}$, $family_{t-1,t-2,t-3}$, and $ftime_{t-1,t-2,t-3}$. The Table also displays the sample units (N) and, in terms of p-values, the Sargan's test for over-identification (Q_S), the Arellano's serial correlation test implicit in the GMM analysis (Q_A), and the Multivariate Ljung-Box Tests (Q_{LB}) for linear dependency among series over time. The significant codes are: *** significance at 1%, ** significance at 5%, and * significance at 10%.

4 Concluding Remarks

In this study, I develop a computational method to improve the existing literature when estimating the effects of obesity, socioeconomic variables, and individual-specific factors on labour market outcomes by dealing with endogeneity problems, causal relationship, and structural model uncertainty. The methodology consists of an econometric model which takes the name of Two-step System Dynamic Panel Data. Firstly, a Bayesian inference is conducted to obtain a subset containing the only (potential) predictors affecting the outcomes. Then, a dynamic longitudinal study is addressed by including all available lags of the variables within the system as instruments to obtain consistent and unbiased estimates.

The application and empirical analysis aim focus on the relationship between high body weight (obesity) and labour market outcomes across Italian regions, by including a set of potential predetermined variables (e.g., lagged values of the variables of interest), endogenous variables (e.g., socieconomic and risk factors affecting labour productivity and social environment), and heterogeneous individual-specific factors possibly correlated with some variables within the system. The data are collected referring to two databases: (i) the Central Institute of Statistics and (ii) the report on equitable and sustainable well-being. The sample units correspond to all the 21 Italian regions and the time period spans the years between 2007 – 2017 in order to cover a sufficiently large sample to address possible causal relationships between obesity, wages, and labour productivity.

By running the first step implicit in the TSDPD model, I find that wage effects are highly affected from a pool of socioeconomic and risk factors (e.g., excess weight, smoking consumption, and alcohol use), a pool of individual-specific heterogeneity (e.g., education level, family relationship, and free time satisfaction), and two predetermined variables accounting for lagged wage effects and lagged employment rate. The latter would be an interesting instrument to investigate how employment prospects affect causal link between wages and obesity, and thus assessed in the TS-DPD model as potential variable of interest.

From a modeling perspective, negative (causal) impacts of excess weight run in both the workplace and wage. The same occurs in the opposite direction: negative health impacts might hold due to excessive amount of time devoted to work or unsatisfactory pay. Socioeconomic factors are negatively related to wage effects and working conditions: highly large smoking and alcohol consumption tend to show negative effects on working conditions and earnings. Both the socioeconomic factors show a highly strong causal relationship with the only labour market productivity and thus potentially affected from other not-directly observed factors in terms of wage penalties (e.g., ability, working hours, and employment status). Wage improvements strongly depend on family relationship satisfaction and even more on free time activities; whereas, positive impacts of good work performance would only depend on lifestyle.

Similar results were found accounting for gender and can be summarised in three main findings. First, obesity wage and employment penalties persist for males and even more for females. Second, socioeconomic factors negatively affect employment opportunities and wage effects for male and slightly more for female individuals. Third, education level, employment opportunities and wage improvements seem to matter more among male individuals.

From a policy perspective, appropriate prevention and health care in support of chronic diseases might lead to consistent gains in economic production through healthier workplace and more active workforce. In a context of prevention policy-relevant strategies, the empirical analysis highlights that adverse labour market outcomes and thus associated production losses depict high costs society in terms of additional cost components. The highest component is related to individuals with highly negative socioeconomic factors and large risk of chronic diseases in terms of labour force. Thus, most of fiscal revenues addressed to public expenditures will be assigned to increase employment opportunities, improve working conditions, and employ welfare spending. It follows that employers will support temporary replacement costs and recurring staff turnover by implying competitive losses in the labour market. The same occurs with other socioeconomic factors such as smoking and alcohol consumption, which are associated with adverse labour market outcomes by involving additional costs for employers as well as workers.

A Summary Results

Table 7 summarises all the results highlighting sign, effect magnitude, and causality by focusing on the only one-way causal link: from TEV^1 to Y^W and from TEV^2 to $P6$.

Table 7: Summary Results

		Labour Market Outcomes	
		`Employment`	`Wage`
Risk and Socioeconomic Factors	`obesity`	Causal Effect (-) Strong Evidence	Causal Effect (-) Strong Evidence
	`smoke`	Causal Effect (-) Strong Evidence	Non-Causal Effect (-) Moderate Evidence
	`alcohol`	Causal Effect (-) Strong Evidence	Non-Causal Effect (-) Moderate Evidence
Heterogeneous Individual Characteristics	`degree`	Causal Effect (+) Strong Evidence	Causal Effect (+) Strong Evidence
	`family`	Non-Causal Effect (+) Moderate Evidence	Non-Causal Effect (+) Moderate Evidence
	`ftime`	Non-Causal Effect (+) Moderate Evidence	Causal Effect (+) Strong Evidence
Predetermined Variables	`L.employ`	–	Causal Effect (+) Strong Evidence
	`L.wage`	Causal Effect (+) Strong Evidence	–

The Table summarises all the results found in Sections 3.2 and 3.3 by focusing on causal link, sign, and effect magnitude between obesity and labour market outcomes, including socioeconomic factors and individual-specific heterogeneity.

Compliance with Ethical Standards

Funding: No funding was used for this study.

Conflict of Interest: The author declares no conflict of interest.

References

Ahn, S. C. and Schmidt, P. (1995). Efficient estimation of models for dynamic panel data. *Journal of Econometrics*, 68(1):5–27.

Anderson, T. W. and Hsiao, C. (1981). Estimation of dynamic models with error components. *Journal of the American Statistical Association*, 76(375):598–606.

Arellano, M. (2003). Panel data econometrics. *Oxford University Press, New York*.

Arellano, M. and Bond, S. (1991). Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58(2):277–297.

Averett, S. and Korenman, S. (1996). The economic reality of the beauty myth. *Journal of Human Resources*, 31:304–330.

Baltagi, B. H. (1995). Econometric analysis of panel data. *Wiley, New York*.

Barnay, T. (2015). Health, work and working conditions: a review of the european economic literature. *European Journal of Health Economics*, 17:693–709.

Baum, C. L. and Ford, W. F. (2004). The wage effects of obesity: a longitudinal study. *Health Economics*, 13:885–899.

Breiman, L. (1992). The little bootstrap and other methods for dimensionality selection in regression: X-fixed prediction error. *Journal of the American Statistical Association*, 87(419):738–754.

Breiman, L. (1995). Better subset regression using the nonnegative garrote. *Technometrics*, 37(4):373–384.

Breiman, L. and Spector, P. (1992). Submodel selection and evaluation in regression. the x-random case. *International Statistical Review*, 60(3):291–319.

Burkhauser, R. V. and Cawley, J. (2008). Beyond bmi: the value of more accurate measures of fatness and obesity in social science research. *Journal of Health Economics*, 27(2):519–529.

Burkhauser, R. V., Cawley, J., and Schmeiser, M. D. (2009). The timing of the rise in u.s. obesity varies with measure of fatness. *Economics and Human Biology*, 7(3):307–318.

Caliendo, M. and Lee, W.-S. (2013). Fat chance! obesity and the transition from unemployment to employment. *Economics and Human Biology*, 11(2):121–133.

Canova, F. and Ciccarelli, M. (2009). Estimating multicountry var models. *International Economic Review*, 50(3):929–959.

Canova, F., Ciccarelli, M., and Ortega, E. (2007). Similarities and convergence in g7 cycles. *Journal of Monetary Economics*, 54(3):850–878.

Cawley, J. (2004). The impact of obesity on wages. *Journal of Human Resources*, 39(2):451–474.

Cawley, J. and Chad, M. (2012). The medical care costs of obesity: An instrumental variables approach. *Journal of Health Economics*, 31:219–230.

Chou, S., Grossman, M., and Saffer, H. (2004). An economic analysis of adult obesity: Results from the behavioral risk factor surveillance system. *Journal of Health Economics*, 23(3):565–587.

Datta, G. N. and Nicolai, K. (2008). Work environment satisfaction and employee health: Panel evidence from denmark, france and spain, 1994-2001. *European Journal of Health Economics*, 9(1):51–61.

Dumitrescu, E. and Hurlin, C. (2012). Testing for granger non-causality in heterogeneous panels. *Economic Modelling*, 29(4):1450–1460.

Flegal, K. M., Carroll, M. D., Kuczmarski, R. J., and Johnson, C. L. (1998). Overweight and obesity in the united states: prevalence and trends, 1960-1994. *International Journal of Obesity and Related Metabolic Disorders*, 22(1):39–47.

Flegal, K. M., Carroll, M. D., Ogden, C., and Curtin, L. R. (2010). Prevalence and trends in obesity among us adults, 1999-2008. *Journal of the American Medical Association*, 303(3):235–241.

Gortmaker, S. L., Must, A., Perrin, J. M., Sobol, A. M., and Dietz, W. H. (1993). Social and economic consequences of overweight in adolescence and young adulthood. *New England Journal of Medicine*, 329:1008–1012.

Harper, B. (2000). Beauty, stature, and the labour market. *Oxford Bulletin of Economics and Statistics*, 62:771–800.

Harris, R. D. and Tzavalis, E. (1999). Inference for unit roots in dynamic panels where the time dimension is fixed. *Journal of Econometrics*, 91(2):201–226.

Holtz-Eakin, D., Newey, W., and Rosen, H. S. (1988). Estimating vector autoregressions with panel data. *Econometrica*, 56(6):1371–1395.

Im, K. S., Pesaran, M., and Shin, Y. (2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics*, 115(1):53–74.

Jarl, J. and Gerdtham, U.-G. (2012). Does drinking affect long-term sickness absence? a sample selection approach correcting for employment and accounting for drinking history. *Applied Economics*, 44(22):2811–2825.

Jusot, F., Khlat, M., Rochereau, T., and Serme, C. (2008). Job loss from poor health, smoking and obesity: a national prospective survey in france. *Journal of Epidemiology & Community Health*, 62(4):332–337.

Kass, R. E. and Raftery, A. E. (1995). Bayes factors. *Journal of American Statistical Association*, 90(430):773–795.

Kline, B. and Tobias, J. L. (2008). The wages of bmi: Bayesian analysis of a skewed treatment-response model with nonparametric endogeneity. *Journal of Applied Econometrics*, 23:767–793.

Komlos, J. and Brabec, M. (2010). The trend of mean bmi values of us adults, birth cohorts 1882-1986 indicates that the obesity epidemic began earlier than hitherto thought. *American Journal of Human Biology*, 22(5):631–638.

Lakdawalla, D., Philipson, T., and Bhattacharya, J. (2005). Welfare-enhancing technological change and the growth of obesity. *American Economic Review*, 95:253–257.

Levin, A., Lin, C.-F., and Chu, C.-S. J. (2002). Unit root tests in panel data: asymptotic and finite-sample properties. *Journal of Econometrics*, 108(1):1–24.

Levine, P. B., Gustafson, T. A., and Velenchik, A. D. (1997). More bad news for smokers? the effects of cigarette smoking on wages. *Industrial and Labor Relations Review*, 50(3):493–509.

Llena-Nozal, A. (2009). The effect of work status and working conditions on mental health in four oecd countries. *National Institute Economic Review*, 209(1):72–87.

Loh, E. (1993). The economic effects of physical appearance. *Social Science Quarterly*, 74(2):420–438.

Lundborg, P., Nystedt, P., and Rooth, D. (2010). No country for fat men? obesity, earnings, skills, and health among 450,000 swedish men. *IZA Discussion Paper*, (4775):1–41.

Madigan, D. and Raftery, A. E. (1994). Model selection and accounting for model uncertainty in graphical models using occam's window. *Journal of American Statistical Association*, 89:1535–1546.

Madigan, D., York, J., and Allard, D. (1995). Bayesian graphical models for discrete data. *International Statistical Review*, 63(2):215–232.

Marchand, A., Parent-Lamarche, A., and Blanc, M. E. (2011). Work and high-risk alcohol consumption in the canadian workforce. *International Journal of Environmental Research and Public Health*, 8:2692–2705.

Miller, A. J. (1984). Selection of subsets of regression variables. *Journal of the Royal Statistical Society: Series A*, 147(3):389–425.

Morris, S. (2007). The impact of obesity on employment. *Labour Economics*, 14(3):413–433.

Mosca, I. (2013). Body mass index, waist circumference and employment: evidence from older irish. *Economics & Human Biology*, 11(4):522–533.

Mullahy, J. and Sindelar, J. L. (1993). Alcoholism, work, and income. *Journal of Labor Economics*, 11(3):494–520.

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica*, 49(6):1417–1426.

Pacifico, A. (2019a). International co-movements and business cycles synchronization across advanced economies: A spbvar evidence. *International Journal of Statistics and Probability*, 8(4):68–85.

Pacifico, A. (2019b). International co-movements and business cycles synchronization across advanced economies: A spbvar evidence. *International Journal of Statistics and Probability*, 8(4):68–85.

Pacifico, A. (2019c). Structural panel bayesian var model to deal with model misspecification and unobserved heterogeneity problems. *Econometrics*, 7(1):1–24.

Pacifico, A. (2020a). Fiscal implications, misspecified dynamics, and international spillover effects across europe: A time-varying multicountry analysis. *forthcoming in International Journal of Statistics and Economics*.

Pacifico, A. (2020b). Robust open bayesian analysis: Overfitting, model uncertainty, and endogeneity issues in multiple regression models. *Econometric Reviews*. Eprint link: `https://doi.org/10.1080/07474938.2020.1770996`.

Pagan, J. and Davila, A. (1997). Obesity, occupational attainment, and earnings. *Social Science Quarterly*, 78(3):756–770.

Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics*, 22(2):265–312.

Pesaran, M. H. and Smith, R. (1995). Estimating long-run relationships from dynamic heterogeneous panels. *Journal of Econometrics*, 68(1):79–113.

Raftery, A. E., Madigan, D., and Hoeting, J. A. (1997). Bayesian model averaging for linear regression models. *Journal of American Statistical Association*, 92(437):179–191.

Raftery, A. E., Madigan, D., and Volinsky, C. T. (1995). Accounting for model uncertainty in survival analysis improves predictive performance. *Bayesian Statistics*, 6:323–349.

Rashad, I., Grossman, M., and Chou, S. (2006). The super size of america: An economic estimation of body mass index and obesity in adults. *Eastern Economic Journal*, 32(1):133–148.

Register, C. A. and Williams, D. R. (1990). Wage effects of obesity among young workers. *Social Science Quarterly*, 71(1):130–141.

Stuckler, D., Basu, S., Suhrcke, M., Coutts, A., and McKee, M. (2009). The public health effect of economic crises and alternative policy responses in europe: an empirical analysis. *The Lancet*, 374(9686):315–323.

Tunceli, K., Li, K., and Williams, L. K. (2006). Long-term effects of obesity on employment and work limitations among u.s. adults, 1986 to 1999. *Obesity (Silver Spring)*, 14(9):1637–1646.