**Report for Statistical Learning and Large Data**

# Non-Standard Work, Unionization and Innovation Capacity

Caterina Manicardi, Stefania Scrofani, Anna Snaidero
*Sant'Anna School of Advanced Studies, Pisa*

## 1 Introduction

From the end of the 70s, after the rise of the Monetarist theory following the stagflation phenomena that broke the empirical regularity found by Phillips (1958) and presented as theoretical proof of the neo-Keynesian theories, market deregulation started to become dominant in political agendas. In particular, unlike in the past, where the decisive instruments were found in the domain of macroeconomic policy, the debate focused on labour market institutions, with the new orthodoxy that views their flexibilisation as key to economic success (Freeman, 2005). Indeed, according to this view, market-oriented policies, deregulation, flexibilisation and decentralisation of wage bargaining would allow a more efficient allocation of resources, boost productivity growth, industrial innovation, and increase employment. Therefore, many economists (Donges, 1985; Giersch, 1985; Davis and Minford, 1986; Gavin, 1986) and the leading international agencies (International Monetary Fund, 1999; Laws and Monitor, 2003; OECD, 1994) pointed at unions and the rigidity of wages and employment as reasons for stagnant economic growth.
These theories have driven decisions for pro-flexibility policies in industrialised nations around the world, which contributed to the progressive weakening of unions and the decline of centralised wage-setting systems (Checchi and Lucifora, 2002; Wallerstein and Western, 2000), leading to increased wage inequality and declining union power (Wallerstein and Western, 2000; Freeman and Medoff, 1984). Consequently, wages have increased on average less than productivity, driving a fall of labour share over GDP documented since the 1980s (Blanchard et al., 1997; Elsby et al., 2013; Karabarbounis and Neiman, 2014; Piketty, 2018).

However, due to the productivity slowdown in recent decades, and particularly after the 2008 global crisis when it stabilised around a growth rate slightly above zero in most countries (Erber et al., 2017), the economic theories opposed to the orthodox consensus have regained importance within the debate around the pro-flexibility policy. The observed phenomena of the global crises, in fact, questioned several assumptions of neoclassical theory to such an extent that some mainstream economists have recognised the importance of demand shocks as a possible trigger mechanism for potential hysteresis behaviour in economies. In this framework, overly flexible labour markets can worsen aggregate demand conditions after crises and recessions, directly impacting productivity growth. Indeed, as theorised by the *Verdoorn-Kaldor law* (Verdoorn, 1949; Kaldor, 1961), the change in aggregate demand affects productivity growth through the endogeneity of investment [1] In addition to

[1] Nicholas Kaldor expanded the Keynesian doctrine by deriving its long-term implications (Camara-Neto and Vernengo, 2012): by recuperating the Hicksian super-multiplier and linking that to the investment accelerator, he

this 'demand channel', the stagnation of real wage leads to a lowering of the labour cost relative to capital, hampering the adoption and development of labour-saving innovations (Labini, 1999; Funk, 2002) through the so-called 'relative cost channel' [2].

Although the productivity slowdown appears to be widespread in most OECD countries, the Italian picture is even more disappointing with an older and deeper slowdown (Dosi et al., 2021b). As shown in (Dosi et al., 2012), which analyse the firm-level dynamics to investigate the flat trend in the Italian aggregate productivity, there is a persistently high degree of heterogeneity along many dimensions of firms' performance, including labour productivity and growth rates, at any level of disaggregation: the support of the sectoral distribution of firms' productivity is wide and does not shrink over time. Besides this general fat-tail property, the left tail of the distribution is much fatter than the right tail. Moreover, their analysis highlights the evident weakness of markets in selecting more efficient firms, which is revealed also by 'evolutionary accounting' exercises showing that the more significant part of productivity growth is generated by the improvement of existing firms' productivity rather than the selection across firms. Indeed, studies on the relationship between efficiency and growth, as the analysis of Bottazzi et al. (2010) on Italian and French firms, show that firm-specific factors typically account for nearly an order of magnitude more than the 'selection' of the variance in firm growth rates (Dosi and Nelson, 2010).
This 'selection channel' too finds its possible cause in the changes in labour market institutions: homogeneous wages, sustained by higher centralisation of wage bargaining, prevent technological laggards from relying on wage-cutting as a survival strategy, leading thus to a higher productivity-driven market selection, which raises the demand for capital goods, increasing their production and boosting technological progress, with the result of a greater average quality and innovative industry dynamics (Antonucci and Pianta, 2002; Kleinknecht, 1998; Hoxha and Kleinknecht, 2020).

The tendency towards an Italian neo-dualism involving the constant presence of a relatively large 'laggard' share in any industry and the particularly harmful Italian productivity slowdown can then be linked to the national reform process aimed at wage and working hours flexibilisation. Indeed, it dates back to 1983-1984 with the introduction of the so-called *work/training contracts*, i.e. fixed-term arrangements for young workers. Ten years later, in 1993, the regulation of productivity-related pay schemes shifted from the national level to the firm/regional level, and in 1997 the so-called *Pacchetto Treu* introduced temporary contracts, part-time employment and apprenticeship schemes. Further liberalisation of part-time, temporary and general non-standard forms of job relationships occurred in 2001 and in 2003 with *Legge Biagi.*
Moreover, despite the rise of doubt towards neoclassical theories assumptions given by the 2008 crisis, in Italy, as in many European countries, the sovereign debt crisis revitalised the debate calling for structural reforms and labour market deregulation to sustain productivity, employment and

made investment endogenous to aggregate demand; moreover, he linked this mechanism with the Verdoorn's law (1949), connecting the change in production volume to variations in labour productivity, so that in the end the growth rate results to be determined by autonomous demand, thus ultimately by wage growth. Given the negative relationship between propensity to consume and income level, this theory also involves significant hypotheses about the role of income distribution in economic growth (Kaldor, 1955), leading to a re-evaluation of wage indexation to productivity as a way to increase the demand growth rate, thus speeding innovation and productivity growth (Kleinknecht and Oostendorp, 2002).

[2]The relative cost channel relies on the theory of 'induced technological change'. Some scholars (Allen, 2009) even argue that this channel was one of the causes that made Britain the starting point of the industrial revolution; indeed, its wages were the highest in the eighteenth century, and this have incentivised the development and adoption of labour-saving machinery (Amatori and Colli, 2017). Similar arguments were developed by Naastepad and Kleinknecht (2004) that explain the Dutch productivity slowdown of recent decades as an effect of the growing age of capital stock induced by the wage moderation policy held by the Wassenaar agreement in 1982.

growth. Coherently, *Legge Fornero* in 2012 and the Jobs Act in 2015 weakened and then abolished the "Articolo 18" of Law 300/1970, i.e. *Statuto dei lavoratori*, that protected workers from invalid dismissals by foreseeing the obligation of reinstatement. Particularly, the Jobs Act introduced a new open-ended contract named *contratto a tutele crescenti* that dispensed the employer from the reintegration of the worker in case of unjust lay-offs and stimulated its diffusion with tax relief to those firms hiring workers with this new contract type. At the same time, the existing threshold for temporary employment set at 20% of a firm's total contracts was removed, resulting in a partial counterbalancing of the incentives to open-ended employment (Fana et al., 2016). Eventually, in 2018 a mild countertrend was provided with the so-called *Decreto dignità* that reduced the maximum length of a temporary contract to two years and yet reinforced the requirements for the employer to extend it after one year.

Such gradual casualisation of work has exacerbated existing cleavages within the working class and created new ones: new hires with *contratto a tutele crescenti* coexist with more protected employees; decentralised wage bargaining fuels pay differentials for equal work paving the way to secondary jobs and part-time ends up being misused to the detriment of full-time employment.

The hostile working environment and the progressive flexibilisation have decreased the average duration of working relationships, undermining the incentive towards the workforce training and favouring trust in labour relations. As Lucidi and Kleinknecht (2010) underline, temporary labour relations imply a short payback period that discourages firms from investing in workers' training, and the uncertainty of the workplace discourages workers from acquiring firm-specific skills since they prefer to invest in general and more marketable skills (Belot et al., 2002). Since the lack of training leads to lower knowledge gathering (Ortega and Marchante, 2010), these mechanisms are particularly harmful to firms marked by Schumpeter mark II technological class, where innovative activity is characterised by 'creative accumulation' and require a long path of collection of tacit specific knowledge.

This 'workforce training and loyalty channel' act mainly boosting the accumulation of knowledge within companies, raising the productivity growth driven by incumbent learning. Moreover, employment protection legislation (EPL) amplifies workers' bargaining power, which can more easily organise themself to request higher wages, with all the aforementioned mechanisms that it entails. Finally, EPL, together with rigid wages, could help prevent hysteresis (Dosi et al., 2018) and boost a productivity-driven market selection of firms, leading to a more innovative and dynamic industrial environment (Dosi et al., 2021a).

Motivated by this ongoing debate, our empirical contribute addresses the relation between innovativeness and workforce composition and degree of unionization for Italian firms in 2015. The rationale was to perform different exercises with the same research question, namely whether unionized firms with standard workforce are more innovative than deunionized ones with a larger share of non-standard workers, to assess the robustness of the result across different methods (and datasets). Section 2 discusses the empirical literature while Section 3 and 4 describe the dataset and methods used for the analysis. Relevant results are presented in Section 5 while future improvements by way of conclusion will follow.

## 2 Literature on empirical evidence

The studies of empirical research in this field reveal handy insights into the aggregate effect of numerical or wage flexibility. They could be divided into two strands based on the level of analysis: the firms' level studies, which aim to capture the effect on firms-level from a single-country perspective, and the cross-country studies, which allow for the analysis of employment protection legislation and the degree of centralization of wage bargaining (CWB).

The first strand of analysis gives significant insights into how the choice of personnel contracts impacts a firm's productivity. The methodology is usually based on panel data econometrics models, where firms' performance proxies are regressed on flexibility indicators, as firms' total turnover or its share of non-standard workers (NSW), or average wages, controlling for possible confounders variables.
The results of this type of study are relatively invariant in stating that workers flexibility is inversely related to labour productivity growth (Lucidi and Kleinknecht, 2010 with Italian data 2001-2003), to the probability of having some R&D and of introducing new products and services (Hoxha and Kleinknecht, 2020 with German data 2007-2015) and to the propensity to introduce both product and process innovation (Cetrulo et al., 2019; Reljic et al.,2021 with Italian data 1994-2016). Flexible wage-setting displays the same adverse effects on process innovation (Wachsen and Blind, 2016), and the average labour costs positively impact the propensity to introduce process innovation and productivity (Reljic et al., 2021; Lucidi and Kleinknecht, 2010). This evidence suggests that NSW's resorting represents a 'survival kit' for technological laggards (Hoxha and Kleinknecht, 2020).
Cirillo and Ricci (2020), after detecting a negative correlation between the share of NSW and both labour productivity and wages, suggest that short-term work contracts may support a vicious cycle of low productivity among minor productive firms. Among the control variables, the firm's size, R&D intensity, industries, or, more generally, the Peneder (2010) classification of those industries are usually utilized, and they assume central importance in capturing the heterogeneity of effects across industries. Indeed, scholars find a more significant negative relationship between flexibility and performance indicators in services and in those industries characterized by high cumulativeness of knowledge and R&D intensity, highlighting how firms with a Schumpeter Mark II innovation model are particularly vulnerable to labour market flexibilization.
Other studies detect a positive impact of workers' councils and employee involvement on firms' performance (Black and Lynch, 2001; Zwick, 2002). They confirm how workers' loyalty and commitment impact firms' productivity and innovativeness. Indeed, as detected by the empirical evidence provided by Svensson (2018), flexible workers display lower trust to standard ones. In confirmation of this hypothesis, Storm and Naastepad (2007) shows that as the degree of flexibility increases, the number of technologies and supervisor control bureaucracies increases.

For what concerns cross-country studies, results are more varied. However, although different studies show a general positive effect of centralized wage bargaining systems and employment protection legislation on productivity and innovativeness, like the ones of Nickell and Layard (1999) on labour productivity and Cetrulo et al. (2019) on the share of firms introducing product and process innovations, no robust studies show an unconditionally negative effect of strict EPL. On the contrary, mainstream scholars underline an adverse effect only in countries with low coordination in industrial relations (Tressel and Scarpetta, 2004; Bassanini and Ernst, 2002).
The theoretical arguments provided by authors are two; the first is that in absence of homogeneous wages and coordination, firms do not have a high expected return from internal training since other firms could poach skilled workers by offering higher wages. The second argument refers to the 'hold-up problem', relevant in decentralized industrial relation systems. In this contest, strict EPL increases workers' bargaining power, worsening this mechanism.
These studies shed light on the heterogeneity of EPL's effects depending on firm's characteristics, finding an adverse effect for small firms, thus confirming the importance of internal flexibility in strict EPL, a channel on which they can less rely. Likewise, the opposite effects on high or low tech, routinized or entrepreneurial companies confirm the harmful effect of worker turnover for knowledge-intensive industries characterized by cumulative innovation paths.
Furthermore, as shown in the meta-analysis of Heimberger (2021), there is a total lack of robust

general empirical evidence of the adverse effects of such rigidities on employment since a zero-mean empirical effect of EPL on unemployment cannot be rejected.

## 3 Dataset overview

The data used for the following analysis are drawn from the survey RIL (*Rilevazione Longitudinale su Imprese e Lavoro*) conducted by INAPP, the Italian National Institute for the Analysis of Public Policies. This survey aims at reporting the movements of firms' labour demand in response to changes in legislation and labour contracts, with some additional information regarding the typology of firms, its innovation level and investments. The analysis that follows is based on the wave of 2015, so referring to information of 2014. Firms in the sample are 30019, representing a population of more than 1 million of firms.

The questionnaire administrated to firms is structured in seven sections: I) Ownership and Management, B) Employees, D) Use of employment contracts, C) Personnel Dynamics and Recruitment, F) Industrial Relationships, L) Innovation, Internationalization and New Markets, H) Credit, Investment and Financial Statements.

In the following, by restricting the scope of the analysis to active or merging firms, we obtained a sample of 29804 firms. We proceeded with the data cleaning and with the imputation of missing data for the turnover using the Loess function: in this way we used the relation between total turnover and size to predict missing values in firms' turnover.

After that we transformed the variables related to workforce, training expenditure and investments in percentage unit, subset the variables of interest and eliminate missing values, so that in the end we remained with 2746 observations of 19 variables listed in Table 1 (5 binary variables for innovation capacity and 14 variables about workforce composition). This database is the one we used for the first part of our analysis while for the second set of methods we relied on a new dataset obtained by grouping firms by their Ateco codes[3], ending up with a total of 1554 observations. In this way binary variables in the first dataset became continuous, representing the percentage of firms with that specific Ateco code with value 1 in the variable, e.g. a value of 60 in the *RD* variable for the first observation (Ateco code 493100) reads as the 60% of firms in that economic activity reported to invest in R&D.

Particularly, the first dataset is used to perform two K-means clusterizations, respectively on workforce variables' principal components and on original variables related to innovation capacity, and to analyze the mapping within the two clusterizations. After that, we did a Linear Discriminant Analysis to classify those firms that pursue R&D and those who do not based on workforce characteristics. We also carried out the same exercise with a quadratic discriminant function to compare LDA and QDA in their predictive capacity. In the second part instead we used the dataset with Ateco observations to do a Canonical Correlation Analysis – always between the workforce variable and the innovation ones – and a comparison between OLS, Ridge, Lasso and Best Subset Selection as models to predict firms propensity to do R&D. A detailed discussion of the methods is provided in the following section.

[3] Italian equivalent to the NACE classification.

| Name | Variable description | Unit |
|---|---|---|
| RD | Share of firms per 6-digits ATECO code that invest in R&D | % |
| ProdInn | Share of firms per 6-digits ATECO code that perform prodict innovation | % |
| ProcInn | Share of firms per 6-digits ATECO code that perform process innovation | % |
| Tech | Share of firms per 6-digits ATECO code that invest in technologies and IT equipment | % |
| Plants | Share of firms per 6-digits ATECO code that invest in plants | % |
| Investments | Share of turnover invested | % |
| Graduate | Share of graduated employees over total number of employees | % |
| Secondary | Share of employees with secondary education over total number of employees | % |
| yo25-34 | Share of employees between 25 and 34 years old over total number of employees | % |
| yo35-49 | Share of employees between 35 and 49 years old over total number of employees | % |
| over50yo | Share of employees over 50 years old over total number of employees | % |
| Executives | Share of executives over total number of employees | % |
| Managers | Share of managers over total number of employees | % |
| Clerks | Share of clerks over total number of employees | % |
| Fixed-term | Share of employees with a fixed-term contract over total number of employees | % |
| Apprenticeship | Share of apprentices over total number of employees | % |
| On_call | Share of employees with on-call contract over total number of employees | % |
| PartTime_Indet | Share of employees with open-end part-time contract over total number of employees | % |
| PartTime_Det | Share of employees with fixed-term part-time contract over total number of employees | % |
| Cocopro | Share of employees with cocopro contract over total number of employees | % |
| Cococo | Share of employees with cocopro contract over total number of employees | % |
| Freelancers | Share of freelancers over total number of employees | % |
| Agents | Share of agents over total number of employees | % |
| Associated | Share of Participation partners over total number of employees | % |
| Trainees | Share of trainees over total number of employees | % |
| NSW | Total share Non Standard Workers over total number of employees | % |
| Unionization | Share of unionized workers over total number of employees | % |
| Strikes | Share of hours not worked per employees in 2014 because of strikes and labour conflicts over the average number of hours worked in that firm per worker | % |
| Fired | Shared of fired employees over total number of employees | % |
| Retired_Ant | Share of early retired over total number of employees | % |
| End_Term | Share of employees at the end of their fixed term contract over total number of employees | % |
| Resigned | Share of resigned over total number of employees | % |
| Hiring | Share of hired employees over total number of employees | % |
| Search | Share of employees the firm is looking over the total number of employees | % |

Table 1: Variables. *Cocopro* and *Cococo* stand for "contratto di collaborazione a progetto" and "contratto di collaborazione coordinata e continuativa". Moreover, *Agents* represent commission-based work and agency contract while Non Standard Workers are defined as Cocopro, Cococo, freelancers, agents, associated and trainees.

# 4 Methods

## 4.1 First dataset (observation = firm)

### *Loess and Cross Validation*

The Loess, LOcally Estimated Scatterplot Smoothing, is one of the most flexible non-parametric regression techniques used to capture general patterns in relationships within two-variable making minimal ex-ante assumptions. The result of a loess application is a line through the moving central tendency of the stressor-response relationship. Loess is essentially used to visually assess the relationship between two variables and is especially useful for large datasets, where trends can be hard to visualize. When we used Lowess in our analysis to impute turnover, we had to specify two relevant parameters: the degree and the span. The first one refers to the degree of relations between the variables analyzed (when the degreee is equal to zero the loess function turns in a weighted moving average, while with two degrees we have a quadratic relation). The most important parameter is however the span, which refers to the length of the interval on which the regression is computed. In particular smaller span allows for lower bias at the expense of higher variance since the regression becomes very sensitive to noise; on the contrary a larger span allows to reduce the variance and the noise at the cost of a higher bias.
To find the optimal span we used cross-validation technique. The cross-validation is a re-sampling technique, which implicate drawing repeated samples from a training set to refit a model of interest on each sample in order to obtain additional information. This process is essential to evaluate a model's performance, also known as model assessment, or to select the proper level of flexibility, i.e. model selection, that is what we did.
Indeed, thanks to this technique we resampled our observations, computed the loess regression with different value of the span, used our functional form to predict the remaining observations and finally assessed the span value that minimizes the prediction error.

### *K-means Clustering*

We also performed a clusterization of our observations based on some variables of interest: we used a type of unsupervised statistical analysis that allowed us to search patterns and discover group structures from data without selecting a response variable. In our research we chose the k-means clusterization rather than the hierarchical one since the former does not require an *ex-ante* definition of the distance function but only a distance matrix (as the hierarchical ones) and the number of clusters, plus it has the advantage of a data-driven technique used to assess the optimal number of clusters. Finally, since the K-means is an iterative algorithm that through a sequential updating of cluster membership and centroid aims to find the clusterization that minimizes the total within-cluster sum of square, it results to be sensitive to the initial condition; for this reason, we reinitialized it one hundred times to find the optimal initial condition, i.e. the one resulting in the highest silhouette width.
In our work, we performed two K-means clustering, one based on the original variables related to innovation capacity and the other on the principal components related to workforce composition, being the number of original variables suitable for dimensionality reduction.

### *Principal Component Analysis*

The principal component analysis is an unsupervised dimension reduction technique used to rep-

resent data in lower dimensions and as a denoise preliminary procedure. In fact, the aim of this analysis is to summarize a large set of variables with a smaller number of representative ones that explain the greatest possible variability of the original set.
Particularly, the principal component directions are directions in feature space along which the original data are highly variable, and such to define subspaces as close as possible to the data cloud. The idea is that if we have $n$ observations that vary in a $p$-dimensional space, maybe not all of these dimensions are equally interesting, and PCA pursues the smallest number of dimensions that allow the observations to vary as much as possible along each of them.
Each of these $z$ PCA dimensions is a normalized linear combination of the p features:

$$z_{i1} = \phi_{11}x_{i1} + \phi_{21}x_{i2} + \; + \phi_{p1}x_{ip}$$

where the $\phi$ are called the loadings, and together they make up the principal component loading vector which defines the direction in feature space along which the data vary the most.

*Linear Discriminant Analysis*

LDA is a classifier used to find a linear transformation of features to classify two or more classes, assuming that such classes are linearly separable and in this way operating a dimensionality reduction. This linear discriminant function would both maximize between-class variance and minimize within-class variance, allowing to project data from a $D$ dimensional feature space down to a $D'(< D)$ dimensional space.
Linear discriminant analysis assumes gaussianity in data and that classes display identical covariance matrices. If these assumptions are not verified, the discriminant function might be a quadratic function, instead of linear, of predictors. In fact, we also performed a Quadratic Discriminant Analysis to compare the predictive ability of both models; to this end, we split data in train and test observations with a 80-20 ratio and defined firms that invest in R&D and those who do not as the two classes.

## 4.2 Second dataset (observation = Ateco code)

*Canonical Correlation Analysis*

Canonical Correlation Analysis is a data reduction technique in a multiple output setting. In fact, given a set of output variables $Y_1, Y_2, ..., Y_k$ to predict from inputs $X_1, X_1, ..., X_p$ and being $U$ and $V$ the sets of uncorrelated linear combinations of the $X$ and $Y$ respectively, each member of $U$ is then paired with a member of $V$ so as to maximize the canonical correlation. The pair $(U_i, V_i)$ is named the $i^{th}$ canonical variate pair or dimension.
For example, consider the first canonical dimension $(U_1, V_1)$ with the two canonical variables $U_1$ and $V_1$ defined as

$$U_1 = a_{11}X_1 + a_{12}X_2 + ... + a_{1p}X_p$$
$$V_1 = b_{11}Y_1 + b_{12}Y_2 + ... + b_{1k}Y_k$$

where coefficients $a$ and $b$ are those that maximize the canonical correlation

$$\rho_1 = \frac{cov(U_1, V_1)}{\sqrt{Var(U_1)Var(V_1)}}$$

Dissimilarly to Principal Component Analysis, which focuses on finding linear combinations that account for the most variance in one dataset, CCA aims at finding linear combinations that account for the most correlation in two datasets (the one with dependent or output variables and the other with input variables).

In what follows, we focused instead on model selection techniques: we wanted to find the correct specification of the model describing the impact of unionisation, labour conflict and workforce composition on investments in R&D. The basic model would take the following form:

$$Y_i = \alpha + \beta X_i + \epsilon_i$$

where $i$ indicates the units of the analysis (in this case 1440 6-digits Ateco code), $Y_i$ is the percentage of firms that have RD investments, and $X_i$ is the set of independent variables defined in Table 1.

### *Variance Inflation Factor*

To detect and remove multicollinearity in an Ordinary Least Squares setting, a preliminary step would be to analyse the correlation matrix within the regressors. However, the correlation matrix detects only the linear correlation between two variables when collinearity problems could arise between three or more variables. To avoid this problem, we assessed multicollinearity using the Variance Inflation Factor (VIF), which is computed as follows:

$$VIF(\hat{\beta}_j) = \frac{1}{1 - R^2_{X_j|X_{-j}}}$$

In practice, the VIF function computes a series of linear regressions where the dependent variable is a different regressor every time while the remaining predictors are included as independent variables. The $R^2$ from each regression is then plugged into the formula of the VIF. The closer the $R^2$ to one, the higher the multicollinearity and the higher the VIF. Following the literature, those predictors displaying a VIF value higher than 5 have been removed.

### *Ridge Regression*

The presence of multicollinearity would increase the Least Square estimator's variance, creating an overfitting problem. Consequently, the estimates obtained through the Least Square would not be trustworthy. We then performed a Ridge Regression to control the multicollinearity problem. In the Ridge regression, the coefficient estimates $\hat{\beta}$ are obtained by minimising the following function:

$$RSS + \lambda \sum_{j=1}^{p} \beta_j^2$$

where $\lambda \geq 0$ is the 'tuning' parameter. When $\lambda = 0$, the estimates of the Ridge regression are the same as the OLS regression, but when $\lambda > 0$, the OLS estimates are corrected with a shrinkage penalty, which shrinks the coefficients to zero without putting any of them exactly to zero. Therefore, the Ridge regression has a different set of estimates for each value of $\lambda$, and the decision of the level of the tuning parameter to choose is critical. We applied cross-validation to decide the level of $\lambda$ that minimises the Root Mean Sum of Squares.

*Lasso Regression*

The Ridge regression overcomes the problem of multicollinearity but does not provide a feature selection of the most important variables for the model, even if it gives some hints. To perform a soft feature selection, we proceeded using a Lasso regression. Using the Lasso, the coefficient estimates $\hat{\beta}$ are obtained by minimising the following function:

$$RSS + \lambda \sum_{j=1}^{p} |\beta_j|$$

where $\lambda \geq 0$ is, as before, the tuning parameter. The Lasso not only shrinks the value of the coefficients towards zero, but it sets some of them precisely to zero. Depending on the value given to $\lambda$, the number of variables retained changes: the higher $\lambda$, the lower the number of features selected. Once again, we used cross-validation to select the value of $\lambda$ that minimises the Root Mean Squared Error.

*Best Subset Selection*

We then compared the results obtained by the soft feature selection with those retained by a hard feature selection, specifically the Best Subset Selection. This method recursively estimates the model using a different subset of the regressors every time, starting from a subset of only one regressor to all the $p$ regressors and taking for every subset all possible permutations of the variables present in the dataset. The Best Subset Selection then compares the $2^p$ different models using different criteria that should estimate the out-of-sample performance of the subsets. Since we would prefer a more parsimonious model, we decided to compare the subsets using the Bayes Information Criterion (BIC); in fact, it has a steeper penalty, favouring smaller models than other criteria.
For Least Square models with $p$ predictors, the BIC takes the form:

$$\frac{1}{n}(RSS + log(n)p\hat{\sigma}^2)$$

where $n$ denotes the sample size, that is the number of variables in the subset under consideration, and $\hat{\sigma}^2$ is the unbiased estimate of the error variance.
Instead of using the BIC or any other method that estimates the out-of-sample performance of the subset of variables under analysis, we can use cross-validation. For this purpose, we split the sample into training and test observations following a 80-20 ratio. First, the model is trained on the train set and then evaluated on the test set, computing the Mean Squared Error. We also performed 10-folds cross-validation where the model is trained recursively on nine folds and then tested on the tenth. Then the out-of-sample Mean Squared Error is computed as the average of the Mean Squared Error computed on the different folds.
To finally select the best model, we compared the Root Mean Squared Error obtained from every model.

# 5 Results

## 5.1 First dataset (observation = firm)

We started by splitting firms' observations database into two: one with all firms' workforce variables and the other with the variables related to their innovative capacity. With the first one, the widest,

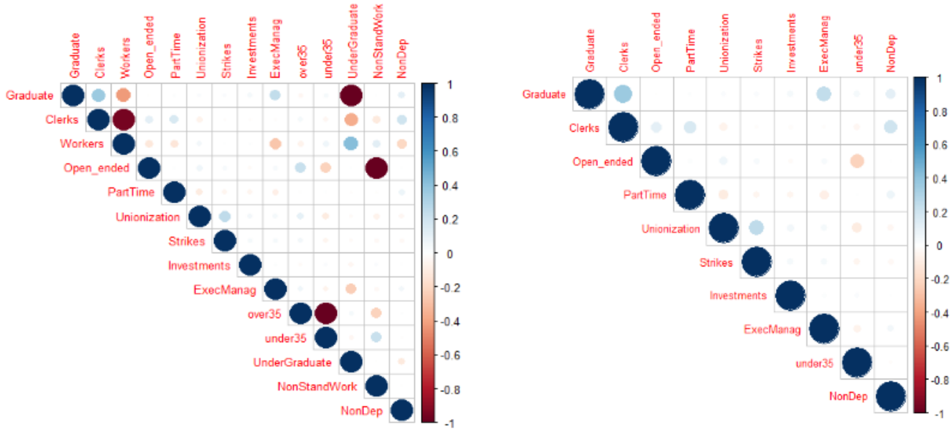



Figure 1: Corrplots

we performed a find correlation function to eliminate the redundant ones, highlighted by dark red circles in Figure 1.

With the selected variables we performed a preliminary PCA, to reduce dimension and smoothing possible noise. Based on the cumulative variance we chose to maintained four components, whose relationships with the variables are illustrated in Figure 2.

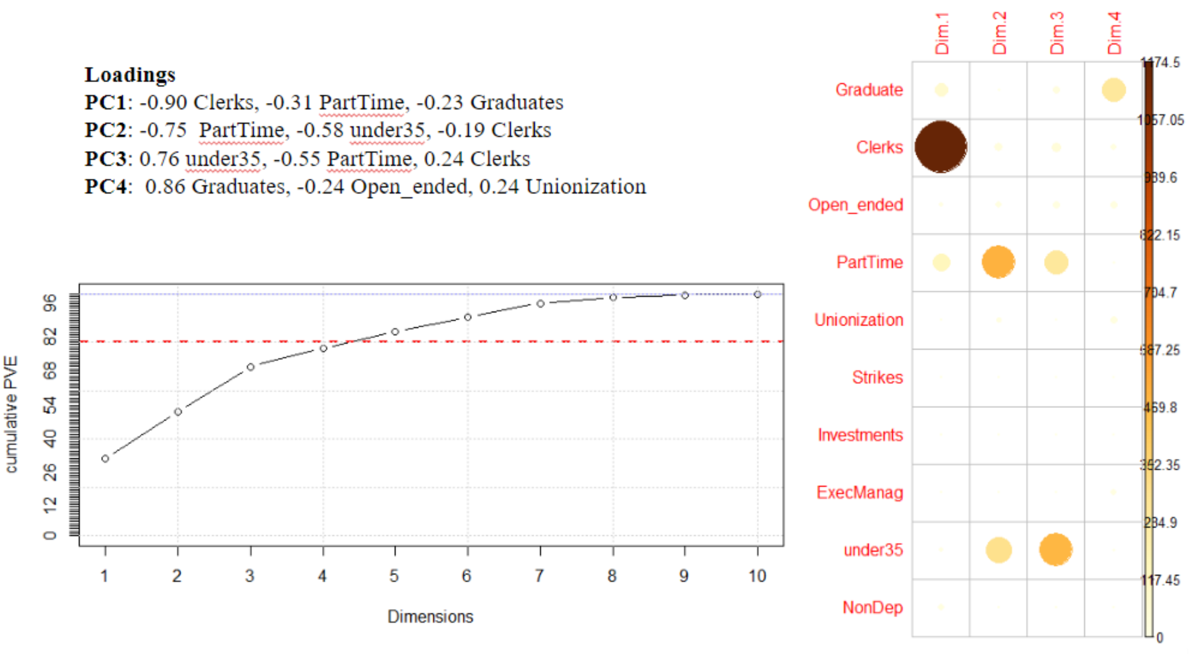



Figure 2: PCA on workforce variables

At this point, the four principal components were used for the K-means clustering. We selected three clusters by looking at the total within sum of square and mapped the clusters on the original variables. As displayed in the table of Figure 3, the average of the three clusters especially diverges in the number of graduates, part-time and unionized workers. Moreover, there is an evident difference in the percentage of clerks, executives, and managers too. To summarize, we can say that the first class is the one characterized by the highest percentage of part-time and non-dependent workers. In contrast, the second class is the most unionized one, with also a high level of strikes and a low percentage of part-time contracts and non dependent workers. Finally, the third can be characterized by its high percentage of graduate workers, clerks, executives and managers.

| | Cluster 1 | Cluster 2 | Cluster 3 |
|---|---|---|---|
| Graduate | 9.10 | 5.09 | 24.54 |
| Clerks | 46.95 | 18.44 | 86.67 |
| Open_ended | 91.98 | 90.09 | 92.59 |
| PartTime | 87.30 | 8.09 | 14.09 |
| NonDep | 10.34 | 6.06 | 13.33 |
| Unionization | 2.25 | 7.71 | 4.85 |
| Strikes | 0.12 | 0.79 | 0.55 |
| Investments | 1.44 | 2.46 | 2.23 |
| ExecManag | 1.64 | 3.36 | 5.31 |
| under35 | 28.84 | 28.71 | 29.93 |

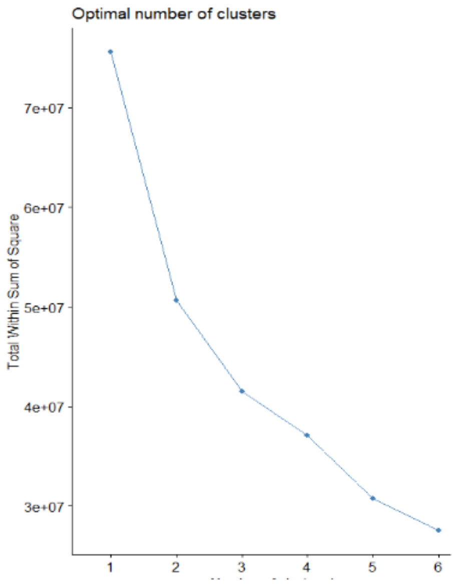



Figure 3: Clusters on workforce variables

The same K-means clustering was then applied based on the variables related to firms' innovative activity. In this case no preliminary PCA was needed since the variables were already few. As we can see from the table in Figure 4, the first cluster is characterize by low values in all the variable of interest (percentage of firms that implement process or product innovation, firms that invest in $R\&D$, in plants or technological equipment), so that we renamed it the group of 'laggard' firms. The second cluster is instead characterized by medium level of propensity to innovate and high level of investment, in particular in technological machinery, so that we referred to this group as the group of 'supply dominated' firms (Pavitt, 1984). Following the Pavitt taxonomy, these are those firms whose innovative capacity comes from the purchase of machinery from specialized companies. Finally, the third group is the one of the innovative firms, indeed it presents very high value in propensity to innovate both products or processes.

| | Cluster 1 | Cluster 2 | Cluster 3 |
|---|---|---|---|
| Product Innovation | 0.24 | 0.33 | 0.83 |
| Process Innovation | 0.16 | 0.22 | 0.81 |
| R&D | 0.04 | 0.26 | 0.31 |
| Plants | 0.18 | 0.63 | 0.60 |
| Tech | 0.09 | 0.74 | 0.37 |

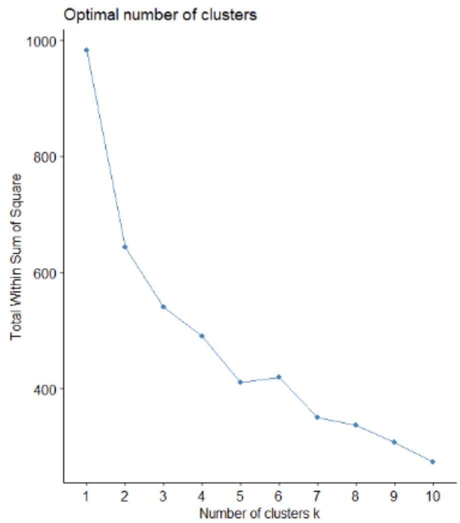



Figure 4: Clusters on innovation variables

Given these two clustering, we looked for a possible correlation between workforce composition and innovative capacity by mapping the percentage of firms belonging to each technological class in each workforce cluster. We found interesting results in line with our starting hypothesis:as shown in Figure 5, the first workforce cluster, characterized by the highest part-time and non-dependent workers, is the one with the highest percentage of 'laggard' firms, while the second, that is the most unionized, displays the highest level of 'innovators'. Finally, the third workforce cluster, with the highest share of graduates and intermediate-upper figures, is highly composed of supply-dominated firms.

| | ClusterWorkForce 1 | ClusterWorkForce 2 | ClusterWorkForce 3 |
|---|---|---|---|
| ClusterTech 1 | 0.57 | 0.33 | 0.28 |
| ClusterTech 2 | 0.24 | 0.29 | 0.46 |
| ClusterTech 3 | 0.19 | 0.39 | 0.27 |

Figure 5: Mapping between the two clusterizations

After that, as an additional exercise we carried out a Linear Discriminant Analysis to classify those firms that invest in research and development. The coefficients of the resulting linear discriminant function are reported in Table 2: it depends negatively on part-time contracts, non standard and not graduated workers, while the highest positive coefficients are those of unionization, strikes and investments. The barplot displayed in Figure 6, which is the density plots of data on the linear discriminant function, confirms our expectations: firms in group 0 ($RD = 0$) have lower values of the function on average compared to firms in group 1 ($RD = 1$), suggesting a higher incidence of non standard workers and part-time contracts, and/or lower unionization and investments levels, in those firms that do not invest in research and development.

| | LD1 |
|---|---|
| Clerks | 0.0023818094 |
| PartTime | -0.0115476254 |
| NonDep | 0.0051126422 |
| Unionization | 0.0273846263 |
| Strikes | 0.1365803508 |
| Investments | 0.0651838205 |
| ExecManag | 0.0175487636 |
| over35 | 0.0018670030 |
| UnderGraduate | -0.0225891643 |
| NonStandWork | -0.0002470122 |

Table 2: Coefficients of the linear discriminant

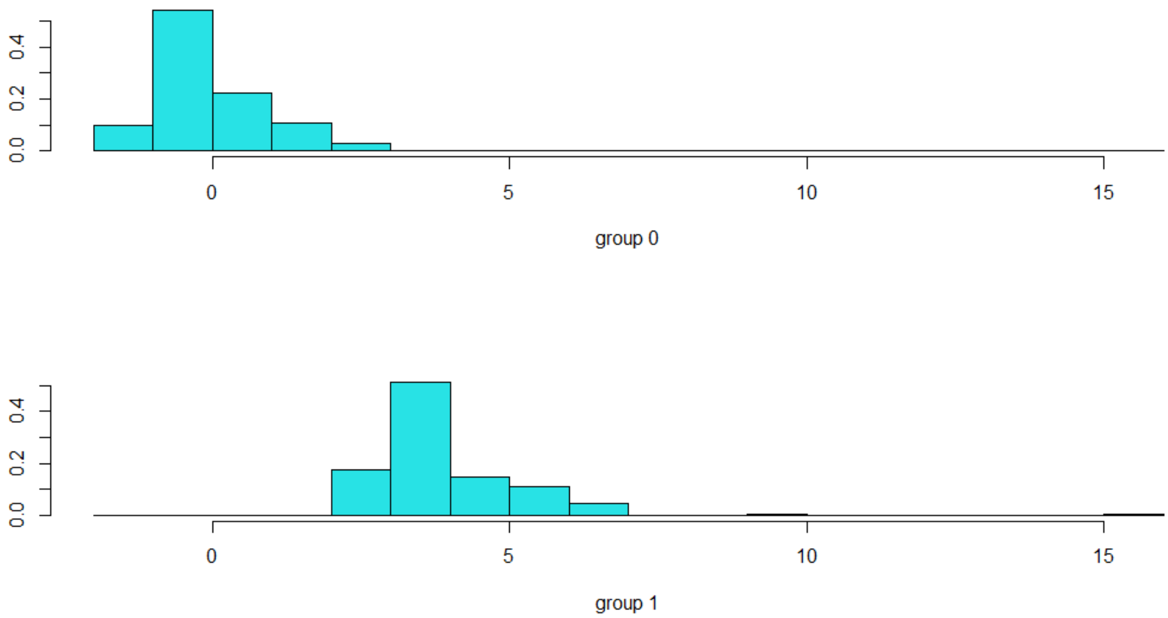
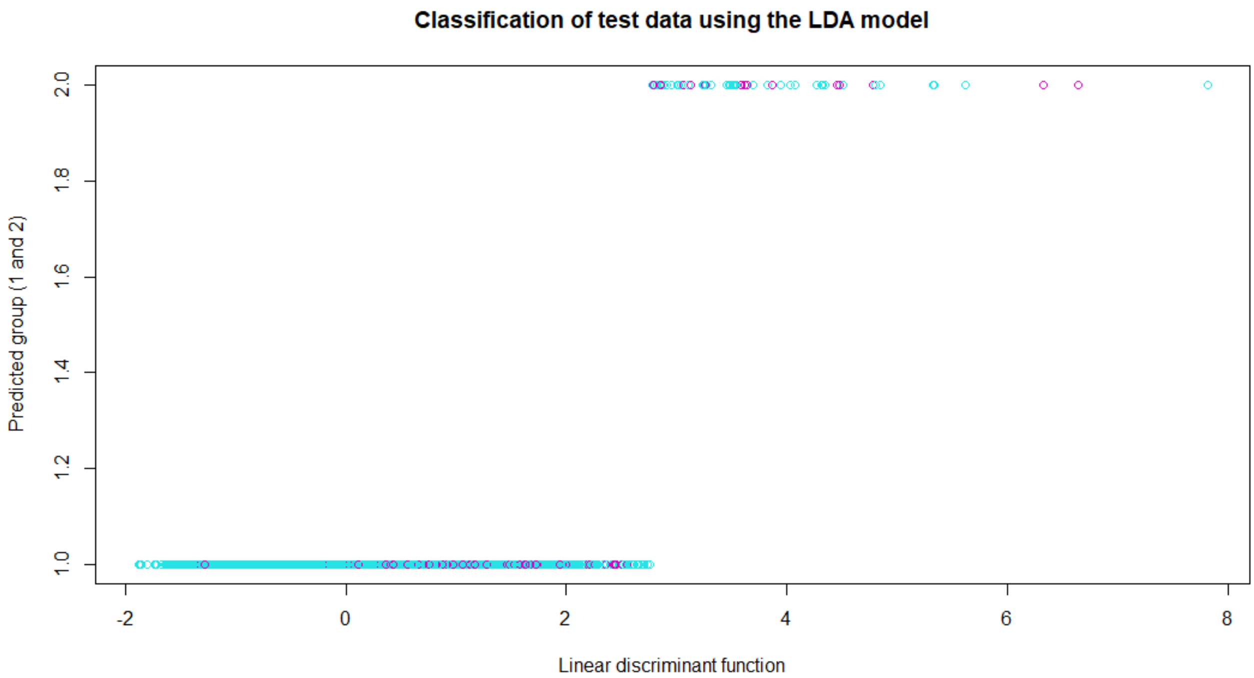



Figure 6: Density plots of data on the linear discriminant function

Since the covariance matrices of the two classes (firms with $RD = 0$ and with $RD = 1$) are different, we also tried to allow for a quadratic discriminant function. Both models provide a relatively high predictive capacity, but the LDA model, despite the violation of the requirement of identical covariance matrices, seems to perform better (see Figure 7).



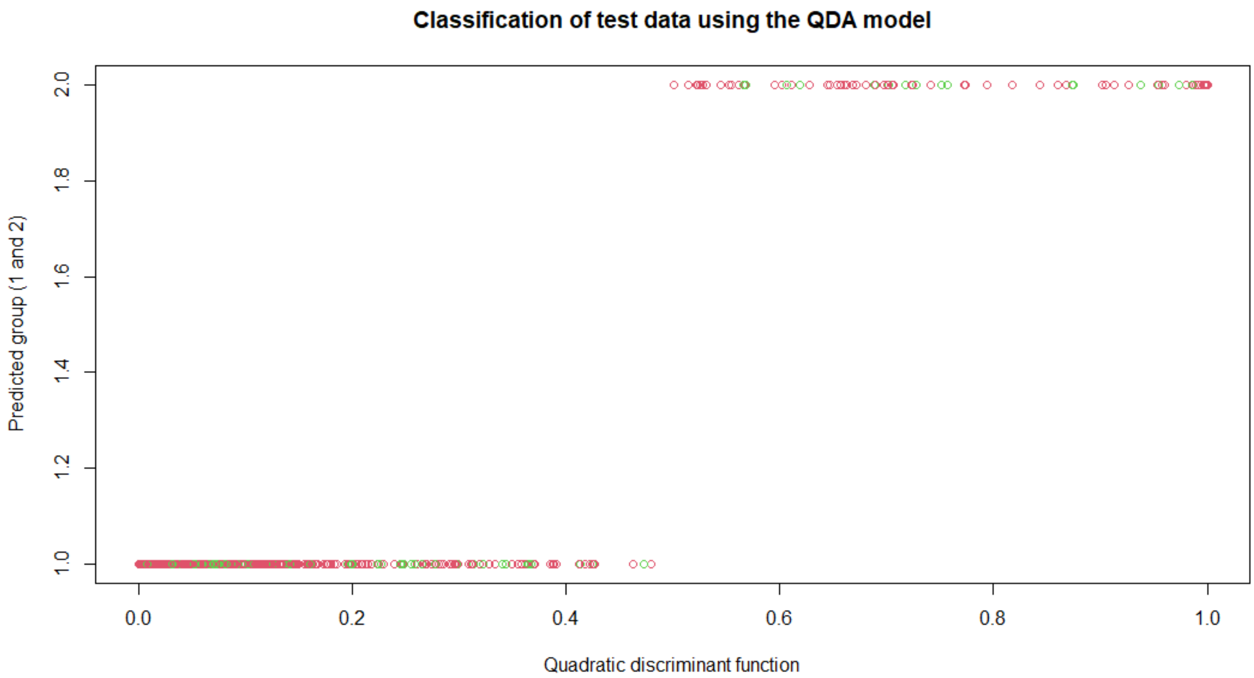




Figure 7: Visual representation of the predictive ability of LDA and QDA models. In this case group 1 and group 2 correspond to $RD = 0$ and $RD = 1$ respectively

## 5.2 Second dataset (observation = Ateco code)

Before proceeding to model selection, the results of cross correlation analysis considering variables related to innovation as the multiple output to predict from the input variables related to workforce are displayed in Figure 8 and in Table 3.
The first canonical component CC1 has a squared canonical correlation of 0.16, which means that 16% of the variation in the first canonical variable of the $Y$ ($CC1\_Y$) is explained by the variation in the first canonical variable of the $X$ ($CC1\_X$). Moreover, $CC1\_X$ depends negatively on unionization and strikes (and agents) while $CC1\_Y$ depends negatively on all original variables related to investments decisions. In the scatter plot of Figure 9 values for each observation of the first two canonical variables is displayed together with a regression line: $CC1\_X$ and $CC1\_Y$ are positively linearly correlated, suggesting that lower unionization degree and labour conflict matches with lower innovation capacity.
The reason for the presence of a threshold of dots at value 0.03 for CC1_Y requires further investigation to exclude misreporting: those observations are economic activities with just one value in innovation-related variables different from zero but the reason for this commonality remains unclear.

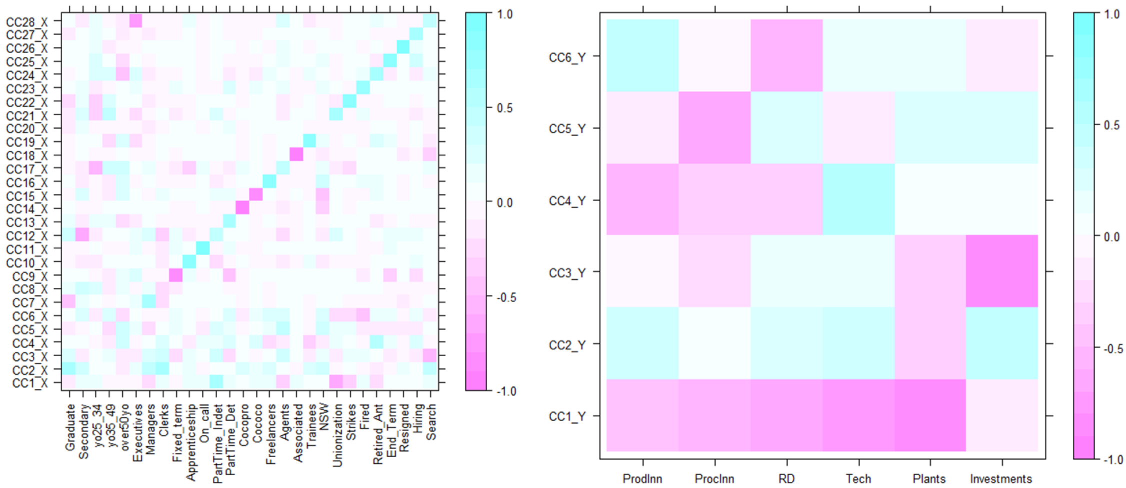


Figure 8: Corrplots for workforce- and innovation-related variables

| | Canonical Correlation | Squared Canonical Correlation |
|---|---|---|
| CC1 | 0.3999 | 0.1599 |
| CC2 | 0.3225 | 0.1040 |
| CC3 | 0.3137 | 0.0984 |
| CC4 | 0.2238 | 0.0501 |
| CC5 | 0.1674 | 0.0280 |
| CC6 | 0.1290 | 0.0167 |

Table 3: Canonical correlation and its squared value for each canonical component

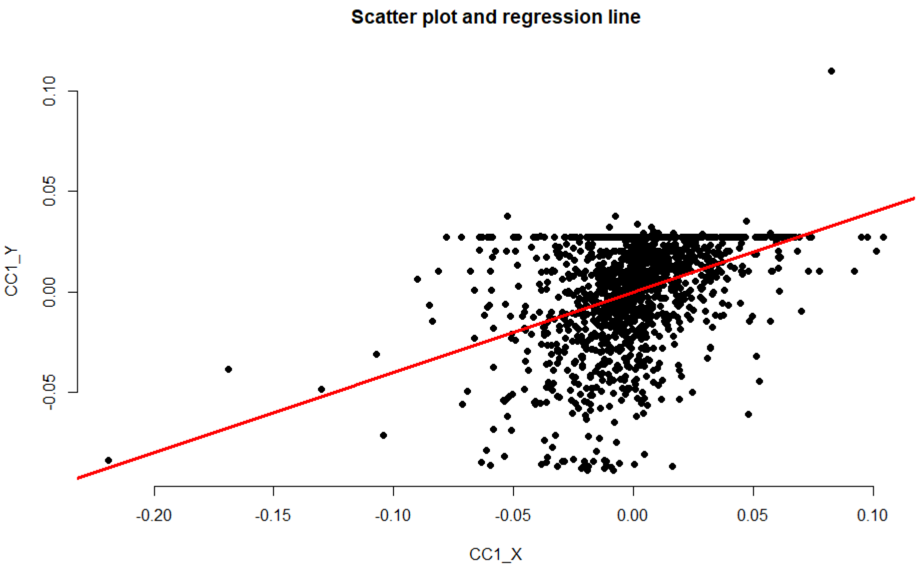



Figure 9: Scatter plot with the first canonical variable for $X$ and $Y$ as coordinates

Overall, the unsupervised analysis and the LDA confirmed the positive relationship between unionization and the percentage of firms that have invested in R&D. Moreover, the CCA has confirmed this positive effect at the 6-digits Ateco code, which is the unit of observation also for the following analysis.

Proceeding with the supervised modelling, we first checked for multicollinearity using the VIF. Figure 10 shows that in our dataset seven variables present severe correlation (VIF$>5$): *yo25_34, yo35_49, over50yo, cocopro, cococo, freelancers, agent and NSW.*

We use the Ridge regression to obtain reliable estimates of the coefficients and p-values of the OLS and overcome the problem of multicollinearity. Figure 11 clearly shows that the percentage of firms that invest in technologies and plants and the percentage of agents in the workforce have a higher effect on the percentage of firms that invest in R&D.

Since we aim to select the best model and verify if the positive effect of the unionization on the R&D is robust, we compare the Ridge regression with models that perform feature selection. The Lasso (Figure 12) still suggests that the most relevant variable for explaining R&D are investments in technologies and plants and the percentage of agents, but unionization is still present in the model that minimizes the MSE. We then compare the Ridge and the Lasso regressions with the Best Subset Selection. Depending on the different criteria, the number and type of feature selected changes. In particular, the Adjusted $R^2$ would suggest retaining seventeen variables, while the Mallow's $C_p$ would retain fourteen. Since we prioritize a parsimonious model, we consider the BIC that retains five variables. As Figure 13 shows, the feature selected includes, once again, unionization.

Table 4 compare the four models used. The coefficient estimates by the Ridge are, in general, slightly smaller than the coefficient estimated by the OLS, but what is even more interesting is that the effects of some regressors on the dependent variable change in the two models. For example, *yo35_49*, *cococo, freeelancers* and *associated* have a positive effect in the OLS regression and a negative effect in the Ridge regression, while *NSW* goes from a negative effect to a positive effect. The regressors that present this change in the effects' direction are exactly those subject to severe correlation in the VIF, pointing to the fact that estimates of an OLS affected by multicollinearity are unreliable. The estimates reported for the Lasso and the Best Subset Selection are the results of an OLS conducted on the variables selected by the two models[4]. The directions of the effects are coherent among the two methods, and the magnitudes of the effects are similar. Unionization is retained as a variable of interest in all the models, and, in all of them, it has a statistically significant positive impact on R&D. These results support the evidence that an increasing flexibilization of work does not stimulate innovation; on the very contrary, a higher percentage of unionized workers is positively related to higher R&D.

Another variable that has a positive and statistically significant effect on R&D in all the models is *Agents*, which calls for further investigation.

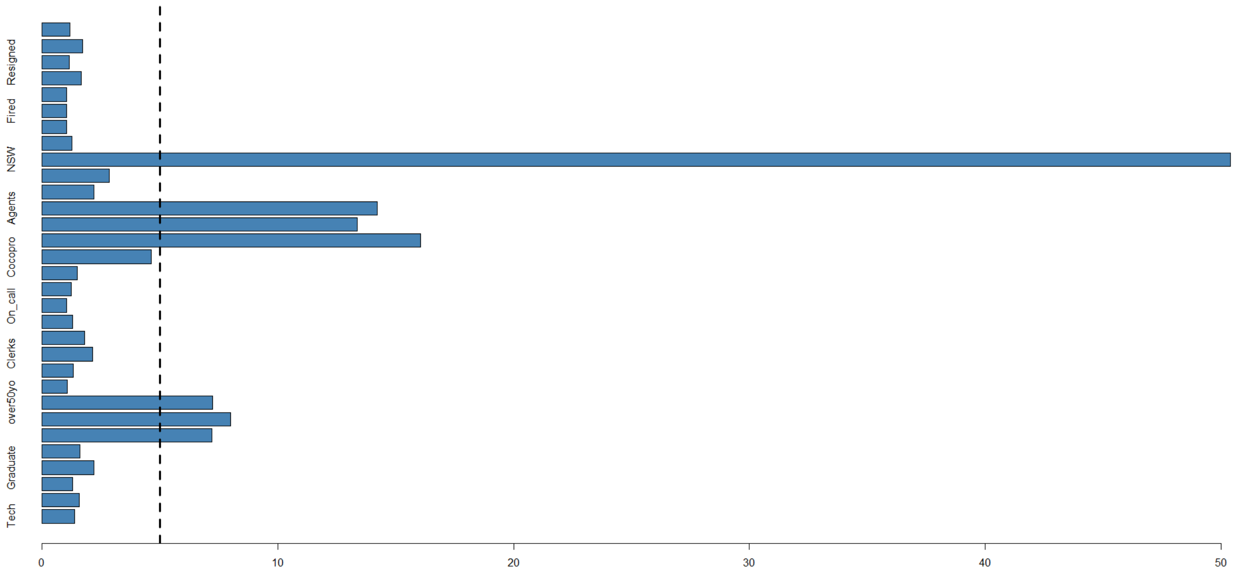



Figure 10: VIF values

[4]For the Best Subset Selection, the estimates obtained through cross-validation are available upon request, but they are very close to the reported OLS estimates.

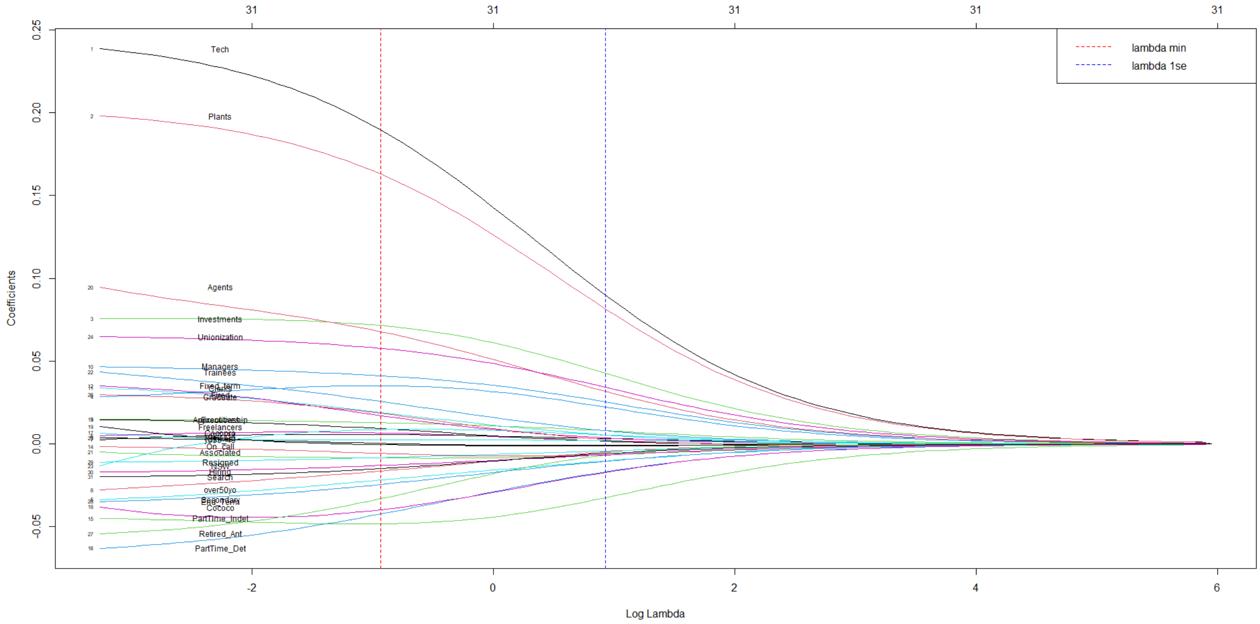



Figure 11: Ridge regression

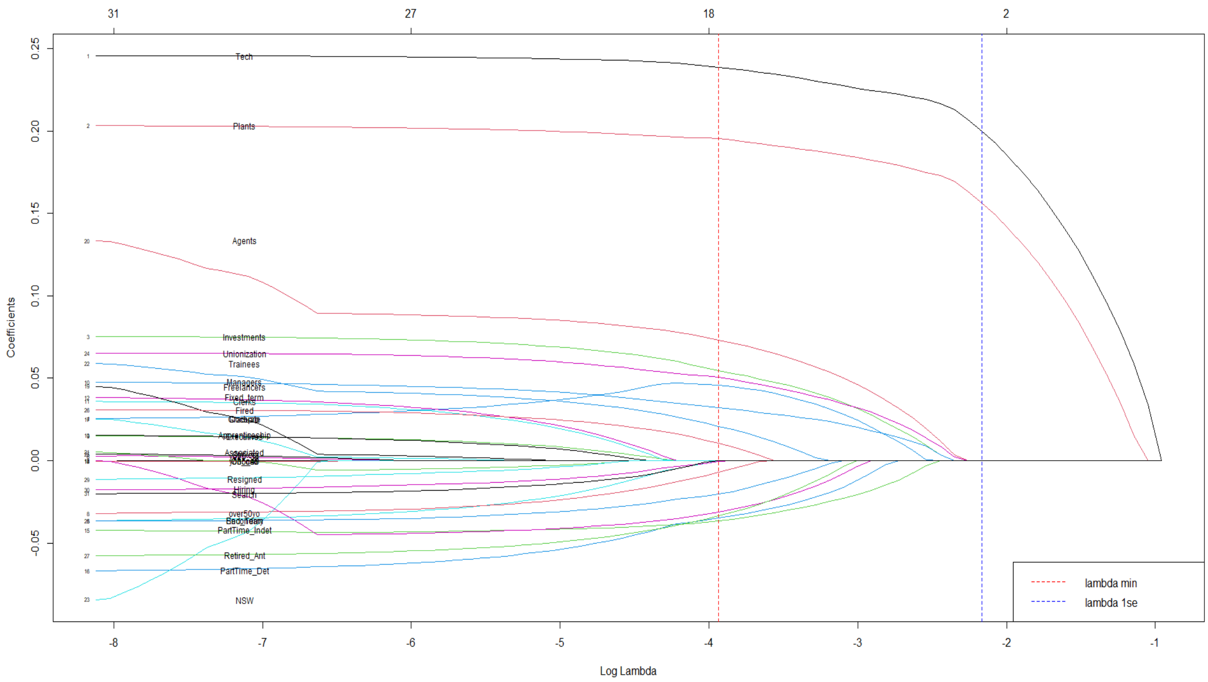



Figure 12: Lasso regression

| | OLS | Ridge | Lasso | Best Subset Selection |
|---|---|---|---|---|
| Tech | 0.201790*** | 1.90E-01 | 0.20398*** | 0.20080*** |
| | (0.022490) | | (0.02220) | (0.02208) |
| Plants | 0.148237*** | 1.63e-01 | 0.14856*** | 0.14995*** |
| | (0.021363) | | (0.02088) | (0.02066) |
| Investments | 0.388324*** | 7.15e-02 | 0.36208*** | 0.33977*** |
| | (0.136498) | | (0.13105) | (0.13082) |
| Graduate | 0.036314 | 3.52e-02 | 0.08205** | 0.11019*** |
| | (0.050684) | | (0.03792) | (0.03482) |
| Secondary | -0.041185 | -2.18e-02 | | |
| | (0.032409) | | | |
| yo25_34 | 0.001235 | 6.70e-03 | | |
| | (0.100469) | | | |
| yo35_49 | -0.005100 | 5.80e-03 | | |
| | (0.095625) | | | |
| over50yo | -0.053139 | -1.64e-02 | -0.04175 | |
| | (0.099559) | | 0.03621 | |
| Executives | 0.071470 | 1.28e-02 | | |
| | (0.114442) | | | |
| Managers | 0.184515* | 4.12e-02 | 0.15837 | |
| | (0.103705) | | (0.10089) | |
| Clerks | 0.033124 | 1.90e-02 | | |
| | (0.031202) | | | |
| Fixed_term | 0.117184 | 1.71e-02 | | |
| | (0.094196) | | | |
| Apprenticeship | 0.069231 | 9.18e-03 | | |
| | (0.118641) | | | |
| On_call | -0.002367 | -5.70e-03 | | |
| | (0.168814) | | | |
| PartTime_Indet | -0.049749 | -4.83e-02 | -0,05645* | |
| | (0.031351) | | (0.02953) | |
| PartTime_Det | -0.416871** | -4.23e-02 | -0.29374* | |
| | (0.179095) | | (0.15004) | |
| Cocopro | 0.288114 | 2.33e-03 | | |
| | (0.381109) | | | |
| Cococo | 0.096375 | -3.97e-02 | -0.17785** | |
| | (0.343330) | | (0.08612) | |
| Freelancers | 0.284035 | -3.77e-04 | | |
| | (0.347968) | | | |
| Agents | 0.616032* | 6.78e-02 | 0.35590*** | 0.38192*** |
| | (0.346346) | | (0.09218) | (0.09234) |
| Associated | 0.173825 | -8.48e-03 | | |
| | (0.473073) | | | |
| Trainees | 0.714253* | 2.57e-02 | 0.42927* | |
| | (0.417187) | | (0.24719) | |
| NSW | -0.264540 | 8.59e-03 | | |
| | (0.334221) | | | |
| Unionization | 0.112489** | 5.77e-02 | 0.10871** | 0.14085*** |
| | (0.044960) | | (0.04382) | (0.04172) |
| Strikes | 0.096221 | 3.83e-04 | | |
| | (0.513495) | | | |
| Fired | 0.137544 | 1.84e-02 | 0.14769 | |
| | (0.106018) | | (0.10417) | |
| Retired_Ant | -2.04448** | -3.31e-02 | -1.96846** | |
| | (0.838225) | | (0.83358) | |
| End_Term | -0.112419 | -2.44e-02 | -0.09667 | |
| | (0.091069) | | (0.08693) | |
| Resigned | -0.052778 | -8.23e-03 | | |
| | (0.118947) | | | |
| Hiring | -0.034442 | -1.30e-02 | -0.02630 | |
| | (0.057835) | | (0.05415) | |
| Search | -0.106785 | -1.50e-02 | | |
| | (0.132875) | | | |
| (Intercept) | 3.9662 | 9.51e-17 | 2.77898 | -0.56993 |
| | (9.4754) | | (1.65881) | (1.02524) |
| Adjusted R squared | 0.2204 | | 0.2253 | 0.2129 |
| RMSE | | 0.8785961 | 0.8775802 | 0.240216 |

Table 4: Model comparison. Robust standard errors in parentheses: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

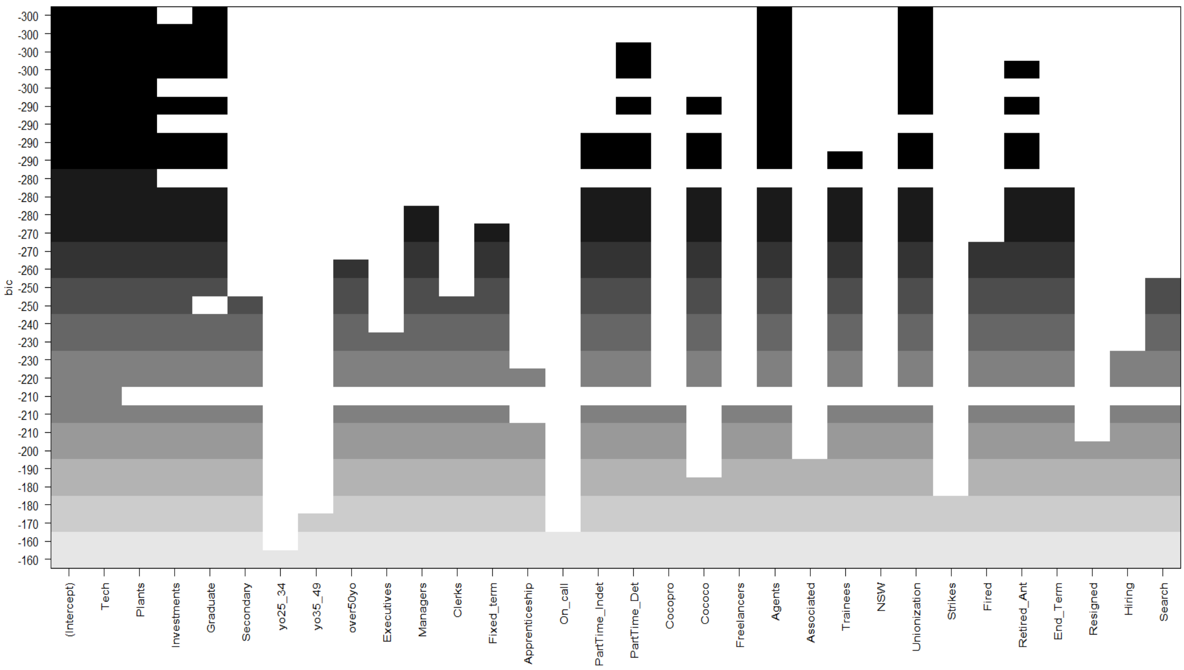



Figure 13: Best Subset Selection

## 6 Further developments by way of conclusion

All the empirical exercises conducted brought evidence that non standard contracts and especially a deunionized workforce have a detrimental effect on firm's innovation capacity. As a further development, it would be interesting to analyse the economic reason why the share of workers with commission-based work and agency contracts seem to play such an important role. Another important ingredient for the economic understanding would concern the spatial dimension of the phenomenon: knowing in which region or province the firm is located would allow for the detection of local or second-level bargaining, as well as for the acknowledgement of regional or province level productive context, i.e. industrial districts, vertical integration and so on. Finally, the section of model selection could be expanded by comparing the four models we analysed using cross-validation and bootstrapping, as well as panel data tracking the firm for multiple years instead of a cross section dataset could help get rid of one-year anomalies (for example in 2014 the Jobs Act was approved, perhaps leading to an overestimation of strikes) and obtain more reliable estimates.

## Bibliography


Allen, R. C. (2009). *The British industrial revolution in global perspective*. Cambridge University Press.

Amatori, F. and Colli, A. (2017). *The Global Economy: A Concise History*. Routledge.

Antonucci, T. and Pianta, M. (2002). Employment effects of product and process innovation in europe. *International Review of Applied Economics*, 16(3):295–307.

Bassanini, A. and Ernst, E. (2002). Labour market institutions, product market regulation, and innovation: cross-country evidence.

Belot, M., Boone, J., and Van Ours, J. C. (2002). Welfare effects of employment protection. *Available at SSRN 319002*.

Black, S. E. and Lynch, L. M. (2001). How to compete: the impact of workplace practices and information technology on productivity. *Review of Economics and statistics*, 83(3):434–445.

Blanchard, O. J., Nordhaus, W. D., and Phelps, E. S. (1997). The medium run. *Brookings papers on economic activity*, 1997(2):89–158.

Bottazzi, G., Dosi, G., Jacoby, N., Secchi, A., and Tamagni, F. (2010). Corporate performances and market selection: some comparative evidence. *Industrial and Corporate Change*, 19(6):1953–1996.

Camara-Neto, A. F. and Vernengo, M. (2012). Keynes after sraffa and kaldor: Effective demand, accumulation and productivity growth. In *Keynes's General Theory*. Edward Elgar Publishing.

Cetrulo, A., Cirillo, V., and Guarascio, D. (2019). Weaker jobs, weaker innovation. exploring the effects of temporary employment on new products. *Applied Economics*, 51(59):6350–6375.

Checchi, D. and Lucifora, C. (2002). Unions and labour market institutions in europe. *Economic Policy*, 17(35):361–408.

Cirillo, V. and Ricci, A. (2020). Heterogeneity matters: temporary employment, productivity and wages in italian firms. *Economia Politica*, pages 1–27.

Davis, J. and Minford, P. (1986). Germany and the european disease (1). *Recherches Économiques de Louvain/Louvain Economic Review*, 52(3-4):373–398.

Donges, J. B. (1985). Chronic unemployment in western europe forever? *World Economy*, 8(4):353–372.

Dosi, G., Freeman, R. B., Pereira, M. C., Roventini, A., and Virgillito, M. E. (2021a). The impact of deunionization on the growth and dispersion of productivity and pay. *Industrial and Corporate Change*, 30(2):377–408.

Dosi, G., Grazzi, M., Tomasi, C., and Zeli, A. (2012). Turbulence underneath the big calm? the micro-evidence behind italian productivity dynamics. *Small Business Economics*, 39(4):1043–1067.

Dosi, G., Guarascio, D., Ricci, A., and Virgillito, M. E. (2021b). Neodualism in the italian business firms: training, organizational capabilities, and productivity distributions. *Small Business Economics*, 57(1):167–189.

Dosi, G. and Nelson, R. R. (2010). Technical change and industrial dynamics as evolutionary processes. *Handbook of the Economics of Innovation*, 1:51–127.

Dosi, G., Pereira, M. C., Roventini, A., and Virgillito, M. E. (2018). Causes and consequences of hysteresis: aggregate demand, productivity, and employment. *Industrial and Corporate Change*, 27(6):1015–1044.

Elsby, M. W., Hobijn, B., and Şahin, A. (2013). The decline of the us labor share. *Brookings Papers on Economic Activity*, 2013(2):1–63.

Erber, G., Fritsche, U., and Harms, P. C. (2017). The global productivity slowdown: Diagnosis, causes and remedies. *Intereconomics*, 52(1):45–50.

Fana, M., Guarascio, D., and Cirillo, V. (2016). Did italy need more labour flexibility? *Intereconomics*, 51(2):79–86.

Freeman, R. B. (2005). Labour market institutions without blinders: The debate over flexibility and labour market performance. *International Economic Journal*, 19(2):129–145.

Freeman, R. B. and Medoff, J. L. (1984). What do unions do. *Indus. & Lab. Rel. Rev.*, 38:244.

Funk, P. (2002). Induced innovation revisited. *Economica*, 69(273):155–171.

Gavin, M. K. (1986). Labor market rigidities and unemployment: the case of severance costs.

Giersch, H. (1985). Eurosclerosis. kiel discussion papers no. 112. *Kie, Germany.*

Heimberger, P. (2021). Does employment protection affect unemployment? a meta-analysis. *Oxford Economic Papers*, 73(3):982–1007.

Hoxha, S. and Kleinknecht, A. (2020). When labour market rigidities are useful for innovation. evidence from german iab firm-level data. *Research Policy*, 49(7):104066.

International Monetary Fund, I. (1999). Chronic unemployment in the euro area: causes and cures. *World Economic Outlook.*

Kaldor, N. (1955). Alternative theories of distribution. *The review of economic studies*, 23(2):83–100.

Kaldor, N. (1961). Capital accumulation and economic growth. In *The theory of capital*, pages 177–222. Springer.

Karabarbounis, L. and Neiman, B. (2014). The global decline of the labor share. *The Quarterly journal of economics*, 129(1):61–103.

Kleinknecht, A. (1998). Is labour market flexibility harmful to innovation? *Cambridge Journal of Economics*, 22(3):387–396.

Kleinknecht, A. and Oostendorp, R. (2002). R&d and export performance: Taking account of simultaneity. In *Innovation and firm performance*, pages 310–320. Springer.

Labini, P. S. (1999). The employment issue: investment, flexibility and the competition of developing countries. *PSL Quarterly Review*, 52(210).

Laws, R. and Monitor, F. (2003). World economic outlook, april 2003: Growth and institutions. *World Economic Outlook.*

Lucidi, F. and Kleinknecht, A. (2010). Little innovation, many jobs: An econometric analysis of the italian labour productivity crisis. *Cambridge Journal of Economics*, 34(3):525–546.

Naastepad, C. W. and Kleinknecht, A. (2004). The dutch productivity slowdown: the culprit at last? *Structural Change and Economic Dynamics*, 15(2):137–163.

Nickell, S. and Layard, R. (1999). Labor market institutions and economic performance. *Handbook of labor economics*, (3):3029–3084.

OECD (1994). *The OECD jobs study: facts, analysis, strategies.* Organisation for Economic Co-operation and Development.

Ortega, B. and Marchante, A. J. (2010). Temporary contracts and labour productivity in spain: a sectoral analysis. *Journal of Productivity analysis*, 34(3):199–212.

Pavitt, K. (1984). Sectoral patterns of technical change: towards a taxonomy and a theory. *Research policy*, 13(6):343–373.

Peneder, M. (2010). Technological regimes and the variety of innovation behaviour: Creating integrated taxonomies of firms and sectors. *Research Policy*, 39(3):323–334.

Piketty, T. (2018). Capital in the twenty-first century. In *Capital in the twenty-first century.* Harvard University Press.

Reljic, J., Cetrulo, A., Cirillo, V., and Coveri, A. (2021). Non-standard work and innovation: evidence from european industries. *Economics of Innovation and New Technology*, pages 1–29.

Storm, S. and Naastepad, C. (2007). *Why labour market regulation may pay off: Worker motivation, co-ordination and productivity growth*, volume 4. International Labour Office.

Svensson, S. (2018). Organizational trust: How to include the division of labour? *Economic and Industrial Democracy*, 39(2):272–293.

Tressel, T. and Scarpetta, S. (2004). *Boosting productivity via innovation and adoption of new technologies: any role for labor market institutions?*, volume 3273. World Bank Publications.

Verdoorn, J. P. (1949). On the factors determining the growth of labor productivity. *Italian economic papers*, 2:59–68.

Wachsen, E. and Blind, K. (2016). More labour market flexibility for more innovation? evidence from employer–employee linked micro data. *Research Policy*, 45(5):941–950.

Wallerstein, M. and Western, B. (2000). Unions in decline? what has changed and why. *Annual Review of Political Science*, 3(1):355–377.

Zwick, T. (2002). Continuous training and firm productivity in germany. Technical report, ZEW Discussion Papers.